



# INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7: C12N 15/54, 9/10, 15/81, 15/82, 1/16, 5/10, A01N 27/067, C12P 7/64

A2

(11) International Publication Number:

WO 00/60095

(43) International Publication Date:

12 October 2000 (12.10.00)

(21) International Application Number:

PCT/EP00/02701

(22) International Filing Date:

28 March 2000 (28.03.00)

(30) Priority Data:

99106656.4 1 April 1999 (01.04.99) EP 99111321.8 10 June 1999 (10.06.99) EP 60/180,687 7 February 2000 (07.02.00) US

(71) Applicant (for all designated States except US): BASF PLANT SCIENCE GMBH [DE/DE]; D-67056 Ludwigshafen (DE).

(72) Inventors; and

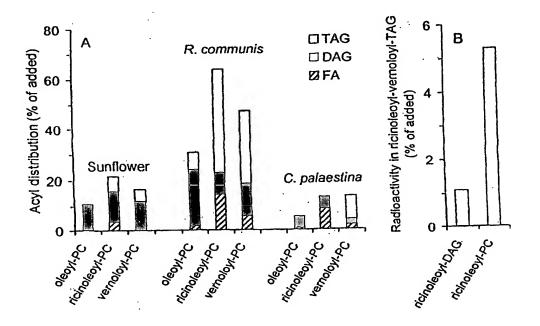
- (75) Inventors/Applicants (for US only): DAHLQVIST, Anders [SE/SE]; Hemmansvägen 2, S-244 66 Furulund (SE). STAHL, Ulf [SE/SE]; Liljegatan 7b, S-753 24 Uppsala (SE). LENMAN, Marit [SE/SE]; Revingegatan 13a, S-223 59 Lund (SE). BANAS, Antoni [SE/PL]; Wiolinowa 14, PL-08110 Siedlce (PL). RONNE, Hans [SE/SE]; Dirigentvägen 169, S-756 54 Uppsala (SE). STYMNE, Sten [SE/SE]; Torrlösa 1380, S-269 90 Svalöv (SE).
- (74) Agent: FITZNER, Uwe; Lintorfer Str. 10, D-40878 Ratingen (DE).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AT, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### **Published**

Without international search report and to be republished upon receipt of that report.

(54) Title: A NEW CLASS OF ENZYMES IN THE BIOSYNTHETIC PATHWAY FOR THE PRODUCTION OF TRIACYLGLYCEROL AND RECOMBINANT DNA MOLECULES ENCODING THESE ENZYMES



#### (57) Abstract

The present invention relates to the isolation, identification and characterization of nucleotide sequences encoding an enzyme catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol, to the said enzymes and a process for the production of triacylglycerols.

## FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

			·				
AL	Albaлia	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
СН		KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DE	•	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		



# A NEW CLASS OF ENZYMES IN THE BIOSYNTHETIC PATHWAY FOR THE PRODUCTION OF TRIACYLGLYCEROL AND RECOMBINANT DNA MOLECULES ENCODING THESE ENZYMES

The present invention relates to the isolation, identification and characterization of recombinant DNA molecules encoding enzymes catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol.

Triacylglycerol (TAG) is the most common lipid-based energy reserve in nature. The main pathway for synthesis of TAG is believed to involve three sequential acyl-transfers from acyl-CoA to a glycerol backbone (1, 2). For many years, acyl-CoA: diacylglycerol acyltransferase (DAGAT), which catalyzes the third acyl transfer reaction, was thought to be the only unique enzyme involved in TAG synthesis. It acts by diverting diacylglycerol (DAG) from membrane lipid synthesis into TAG (2). Genes encoding this enzyme were recently identified both in the mouse (3) and in plants (4, 5), and the encoded proteins were shown to be homologous to acyl-CoA: cholesterol acyltransferase (ACAT). It was also recently reported that another DAGAT exists in the oleaginous fungus *Mortierella ramanniana*, which is unrelated to the mouse DAGAT, the ACAT gene family or to any other known gene (6).

The instant invention relates to novel type of enzymes and their encoding genes for transformation. More specifically, the invention relates to use of a type of genes encoding a not previously described type of enzymes hereinafter designated phospholipid:diacylglycerol acyltransferases (PDAT), whereby this enzyme catalyses an acyl-CoA-independent reaction. The said type of genes expressed alone in transgenic organisms will enhance the total amount of oil (triacylglycerols) produced in the cells. The PDAT genes, in combination with a gene for the synthesis of an uncommon fatty acid will, when expressed in transgenic organisms, enhance the levels of the uncommon fatty acids in the triacylglycerols.

25



15

20

25

30

2



There is considerable interest world-wide in producing chemical feedstock, such as fatty acids, for industrial use from renewable plant resources rather than non-renewable petrochemicals. This concept has broad appeal to manufacturers and consumers on the basis of resource conservation and provides significant opportunity to develop new industrial crops for agriculture.

There is a diverse array of unusual fatty acids in oils from wild plant species and these have been well characterised. Many of these acids have industrial potential and this has led to interest in domesticating relevant plant species to enable agricultural production of particular fatty acids.

Development in genetic engineering technologies combined with greater understanding of the biosynthesis of unusual fatty acids now makes it possible to transfer genes coding for key enzymes involved in the synthesis of a particular fatty acid from a wild species into domesticated oilseed crops. In this way individual fatty acids can be produced in high purity and quantities at moderate costs.

In all crops like rape, sunflower, oilpalm etc., the oil (i.e. triacylglycerols) is the most valuable product of the seeds or fruits and other compounds like starch, protein, and fibre is regarded as by-products with less value. Enhancing the quantity of oil per weight basis at the expense of other compounds in oil crops would therefore increase the value of crop. If genes, regulating the allocation of reduced carbon into the production of oil can be up-regulated, the cells will accumulate more oil on the expense of other products. Such genes might not only be used in already high oil producing cells, such as oil crops, but could also induce significant oil production in moderate or low oil containing crops such as e.g. soy, oat, maize, potato, sugarbeats, and turnips as well as in micro-organisms.



## Summary of the invention

Many of the unusual fatty acids of interest, e.g. medium chain fatty acids, hydroxy fatty acids, epoxy fatty acids and acetylenic fatty acids, have physical properties that are distinctly different from the common plant fatty acids. The present inventors have found that, in plant species naturally accumulating these uncommon fatty acids in their seed oil (i.e. triacylglycerol), these acids are absent, or present in very low amounts in the membrane (phospho)lipids of the seed. The low concentration of these acids in the membrane lipids is most likely a prerequisite for proper membrane function and thereby for proper cell functions. One aspect of the invention is that seeds of transgenic crops can be made to accumulate high amounts of uncommon fatty acids if these fatty acids are efficiently removed from the membrane lipids and channelled into seed triacylglycerols.

15

20

25

30

The inventors have identified a novel class of enzymes in plants catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the production of triacylglycerol through an acyl-CoA-independent reaction and that these enzymes (phospholipid:diacylglycerol acyltransferases, abbreviated as PDAT) are involved in the removal of hydroxylated, epoxygenated fatty acids, and probably also other uncommon fatty acids such as medium chain fatty acids, from phospholipids in plants.

This enzyme reaction was shown to be present in microsomal preparations from baker's yeast (*Saccharomyces cerevisiae*). The instant invention further pertains to an enzyme comprising an amino acid sequence as set forth in SEQ ID No. 2 or a functional fragment, derivate, allele, homolog or isoenzyme thereof. A so called ,knock out' yeast mutant, disrupted in the respective gene was obtained and microsomal membranes from the mutant was shown to totally lack PDAT activity. Thus, it was proved that the disrupted gene encodes a PDAT enzyme (SEQ ID NO. 1 and 2). Furtherm, this PDAT enzyme is

15

20

25

30



characterized through the amino acid sequence as set forth in SEQ ID NO 2 containing a lipase motif of the conserved sequence string FXKWVEA.

The instant invention pertains further to an enzyme comprising an amino acid sequence as set forth in SEQ ID NO. 1a, 2b or 5a or a functional fragment, derivate, allele, homolog or isoenzyme thereof.

Further genes and/or proteins of so far unknown function were identified and are contemplated within the scope of the instant invention. A gene from Schizosaccharomyces pombe, SPBC776.14 (SEQ ID. NO. 3), a putative open reading frame CAA22887 of the SPBC776.14 (SEQ ID NO. 13) were identified.

Further Arabidopsis thaliana genomic sequences (SEQ ID NO. 4, 10 and 11) coding for putative proteins were identified, as well as a putative open reading frame AAC80628 from the A. thaliana locus AC 004557 (SEQ ID NO. 14) and a putative open reading frame AAD10668 from the A. thaliana locus AC 003027 (SEQ ID NO. 15) were identified.

Also, a partially sequenced cDNA clone from Neurospora crassa (SEQ ID NO. 9) and a Zea mays EST (Extended Sequence Tac) clone (SEQ ID NO. 7) and corresponding putative amino acid sequence (SEQ ID NO. 8) were identified. Finally, two cDNA clones were identified, one Arabidopsis thaliana EST (SEQ ID NO. 5 and corresponding predicted amino acid sequence SEQ ID NO. 6) and a Lycopersicon esculentum EST clone (SEQ ID NO. 12) were identified. Further, enzymes designated as PDAT comprising an amino acid sequence selected from the group consisting of sequences as set forth in SEQ ID NO 2a, 3a, 5b, 6 or 7b containing a lipase motif FXKWVEA are contemplated within the scope of the invention. Moreover, an enzyme comprising an amino acid sequence encoded through a nucleotide sequence, a portion, derivate, allele or homolog thereof selected from the group consisting of sequences as set forth in SEQ ID No. 1, 1b, 3, 3b, 4, 4a, 4b, 5, 5b, 6b, 7, 8b, 9, 9b, 10, 10b, 11, 11b or 12 or a functional fragment, derivate, allele, homolog or isoenzyme of the enzyme encoding amino acid sequence are included within the scope of the invention.

10

15

20

25

30

A functional fragment of the instant enzyme is understood to be any polypeptide sequence which shows specific enzyme activity of a phospholipid:diacylglycerol acyltransferase (PDAT). The length of the functional fragment can for example vary in a range from about  $660 \pm 10$  amino acids to  $660 \pm 250$  amino acids, preferably from about  $660 \pm 50$  to  $660 \pm 100$  amino acids, whereby the "basic number" of 660 amino acids corresponds in this case to the polypeptide chain of the PDAT enzyme of SEQ ID NO. 2 encoded by a nucleotide sequence according to SEQ ID NO. 1. Consequently, the "basic number" of functional fullength enzyme can vary in correspondance to the encoding nucleotide sequence.

A portion of the instant nucleotide sequence is meant to be any nucleotide sequence encoding a polypeptid which shows specific activity of a phospholipid:diacylglycerol acyltransferase (PDAT). The length of the nucleotide portion can vary in a wide range of about several hundreds of nucleotides based upon the coding region of the gene or a highly conserved sequence. For example the length varies in a range form about  $1900 \pm 10$  to  $1900 \pm 1000$  nucleotides, preferably form about  $1900 \pm 50$  to  $1900 \pm 700$  and more preferably form about  $1900 \pm 100$  to  $1900 \pm 500$  nucleotides, whereby the "basic number" of 1900 nucleotides corresponds in this case to the encoding nucleotide sequence of the PDAT enzyme of SEQ ID NO. 1. Consequently, the "basic number" of functional fullength gene can vary.

An allelic variant of the instant nucleotide sequence is understood to be any different nucleotide sequence which encodes a polypeptide with a functionally equivalent function. The alleles pertain naturally occuring variants of the instant nucleotide sequences as well as synthetic nucleotide sequences produced by methods known in the art. Contemplated are even altered nucleotide sequences which result in an enzyme with altered activity and/or regulation or which is resistant against specific inhibitors. The instant invention further includes natural or synthetic mutations of the originally isolated nucleotide

PCT/EP00/02701

sequences. These mutations can be substitution, addition, deletion, inversion or insertion of one or more nucleotides.

6

A homologous nucleotide sequence is understood to be a complementary sequence and/or a sequence which specifically hybridizes with the instant nucleotide sequence. Hybridizing sequences include similar sequences selected from the group of DNA or RNA which specifically interact to the instant nucleotide sequences under at least moderate stringency conditions which are known in the art. A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. This further includes short nucleotide sequences of e.g. 10 to 30 nucleotides, preferably 12 to 15 nucleotides. Included are also primer or hybridization probes.

15

20

25

30

10

A homologous nucleotide sequence included within the scope of the instant invention is a sequence which is at least about 40%, preferably at least about 50 % or 60%, and more preferably at least about 70%, 80% or 90% and most preferably at least about 95%, 96%, 97%, 98% or 99% or more homologous to a nucleotide sequence of SEQ ID NO. 1.

All of the aforementioned definitions are true for amino acid sequences and functional enzymes and can easily transferred by a person skilled in the art.

Isoenzymes are understood to be enzymes which have the same or a similar substrate specifity and/or catalytic activity but a different primary structure.

In a first embodiment, this invention is directed to nucleic acid sequences that encode a PDAT. This includes sequences that encode biologically active PDATs as well as sequences that are to be used as probes, vectors for transformation or cloning intermediates. The PDAT encoding sequence may



encode a complete or partial sequence depending upon the intended use. All or a portion of the genomic sequence, cDNA sequence, precursor PDAT or mature PDAT is intended.

Further included is a nucleotide sequence selected from the group consisting of sequences set forth in SEQ ID No. 1, 1b, 3, 3b, 4, 4a, 4b, 9b, 10, 10b or 11 or a portion, derivate, allele or homolog thereof. The invention pertains a partial nucleotide sequence corresponding to a fullength nucleotide sequence selected from the group consisting of sequences set forth in SEQ ID No. 5, 5b, 6b, 7, 8b, 9, 11b or 12 or a portion, derivate, allele or homolog thereof. Moreover, a nucleotide sequence comprising a nucleotide sequence which is at least 40% homologous to a nucleotide sequence selected form the group consisting of those sequences set forth in SEQ ID No. 1 1b, 3, 3b, 4, 4a, 4b, 5, 5b, 6b, 7, 8b, 9, 9b, 10, 10b, 11, 11b or 12 is contemplated within the scope of the invention.

15

20

25

30

The instant invention pertains to a gene construct comprising a said nucleotide sequences of the instant invention which is operably linked to a heterologous nucleic acid.

The term operably linked means a serial organisation e.g. of a promotor, coding sequence, terminator and/or further regulatory elements whereby each element can fulfill its original function during expression of the nucleotide sequence.

Further, a vector comprising of a said nucleotide sequence of the instant invention is contemplated in the instant invention. This includes also an expression vector as well as a vector further comprising a selectable marker gene and/or nucleotide sequences for the replication in a host cell and/or the integration into the genome of the host cell.

In a different aspect, this invention relates to a method for producing a PDAT in a host cell or progeny thereof, including genetically engineered oil seeds, yeast and moulds or any other oil accumulating organism, via the expression of a



construct in the cell. Cells containing a PDAT as a result of the production of the PDAT encoding sequence are also contemplated within the scope of the invention.

Further, the invention pertains a transgenic cell or organism containing a said nucleotide sequence and/or a said gene construct and/or a said vector. The object of the instant invention is further a transgenic cell or organism which is an eucaryotic cell or organism. Preferably, the transgenic cell or organism is a yeast cell or a plant cell or a plant. The instant invention further pertains said transgenic cell or organism having an altered biosynthetic pathway for the production of triacylglycerol. A transgenic cell or organism having an altered oil content is also contemplated within the scope of this invention.

Further, the invention pertains a transgenic cell or organism wherein the activity of PDAT is altered in said cell or organism. This altered activity of PDAT is characterized by an alteration in gene expression, catalytic activity and/or regulation of activity of the enzyme. Moreover, a transgenic cell or organism is included in the instant invention, wherein the altered biosynthetic pathway for the production of triacylglycerol is characterized by the prevention of accumulation of undesirable fatty acids in the membrane lipids.

In a different embodiment, this invention also relates to methods of using a DNA sequence encoding a PDAT for increasing the oil-content within a cell.

Another aspect of the invention relates to the accommodation of high amounts of uncomman fatty acids in the triacylglycerol produced within a cell, by introducing a DNA sequence producing a PDAT that specifically removes these fatty acids from the membrane lipids of the cell and channel them into triacylglycerol. Plant cells having such a modification are also contemplated herein.

15

Further, the invention pertains a process for the production of triacylglycerol, comprising growing a said transgenic cell or organism under conditions whereby the said nucleotide sequence is expressed and whereby the said transgenic cells comprising a said enzyme catalysing the transfer of fatty acids from phospholipids to diacylglycerol forming triacylglycerol.

Moreover, triacylglycerols produced by the aforementioned process are included in scope of the instant invention.

Object of the instant invention is further the use of an instant nucleotide sequence and/or a said enzyme for the production of triacylglycerol and/or triacylglycerols with uncommon fatty acids. The use of a said instant nucleotide sequence and/or a said enzyme of the instant invention for the transformation of any cell or organism in order to be expressed in this cell or organism and result in an altered, preferably increased oil content of this cell or organism is also contemplated within the scope of the instant invention.

A PDAT of this invention includes any sequence of amino acids, such as a protein, polypeptide or peptide fragment obtainable from a microorganism, animal or plant source that demonstrates the ability to catalyse the production of triacylglycerol from a phospholipid and diacylglycerol under enzyme reactive conditions. By "enzyme reactive conditions" is meant that any necessary conditions are available in an environment (e.g., such factors as temperature, pH, lack of inhibiting substances) which will permit the enzyme to function.

25

30

20

5

Other PDATs are obtainable from the specific sequences provided herein. Furthermore, it will be apparent that one can obtain natural and synthetic PDATs, including modified amino acid sequences and starting materials for synthetic-protein modelling from the examplified PDATs and from PDATs which are obtained through the use of such examplified sequences. Modified amino acid sequences include sequences that have been mutated, truncated.

10

15

25

30



increased and the like, whether such sequences were partially or wholly synthesised. Sequences that are actually purified from plant preparations or are identical or encode identical proteins thereto, regardless of the method used to obtain the protein or sequence, are equally considered naturally derived.

Further, the nucleic acid probes (DNA and RNA) of the present invention can be used to screen and recover "homologous" or "related" PDATs from a variety of plant and microbial sources.

Further, it is also apparent that a person skilled in the art can, with the information provided in this application, in any organism identify a PDAT activity, purify an enzyme with this activity and thereby identify a "non-homologous" nucleic acid sequence encoding such an enzyme.

The present invention can be essentially characterized by the following aspects:

- 1. Use of a PDAT gene (genomic clone or cDNA) for transformation.
- 20 2. Use of a DNA molecule according to item 1 wherein said DNA is used for transformation of any organism in order to be expressed in this organism and result in an active recombinant PDAT enzyme in order to increase oil content of the organism.
  - 3. Use of a DNA molecule of item 1 wherein said DNA is used for transformation of any organism in order to prevent the accumulation of undesirable fatty acids in the membrane lipids.
  - 4. Use according to item 1, wherein said PDAT gene is used for transforming transgenic oil accumulating organisms engineered to produce any uncommon fatty acid which is harmful if present in high amounts in membrane lipids, such as medium chain fatty acids, hydroxylated fatty acids, epoxygenated fatty acids and acetylenic fatty acids.

10

15

- 5. Use according to item 1, wherein said PDAT gene is used for transforming organisms, and wherein said organisms are crossed with other oil accumulating organisms engineered to produce any uncommon fatty acid which is harmful if present in high amounts in membrane lipids, comprising medium chain fatty acids, hydroxylated fatty acids, epoxygenated fatty acids and acetylenic fatty acids.
- 6. Use according to item 1, wherein the enzyme encoded by said PDAT gene or cDNA is coding for a PDAT with distinct acyl specificity.
- 7. Use according to item 1 wherein said PDAT encoding gene or cDNA, is derived from *Saccharomyces cereviseae*, or contain nucleotide sequences coding for an amino acid sequence 30% or more identical to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
  - 8. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from *Saccharornyces cereviseae*, or contain nucleotide sequences coding for an amino acid sequence 40% or more *identical* to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
  - 9. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from *Saccharomyces cereviseae*, or contain nucleotide sequences coding for an amino acid sequence 60% or more *identical* to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
  - 10. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from *Saccharornyces cereviseae*, or contain nucleotide sequences coding for an amino acid sequence 80% or more identical to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
- 25 11. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from plants or contain nucleotide sequences coding for an amino acid sequence 40% or more identical to the amino acid sequence of PDAT from *Arabidopsis thaliana* or to the protein encoded by the fullength counterpart of the partial Zea mays, Lycopericon esculentum, or Neurospora crassa cDNA clones.

15

20

:



12. Transgenic oil accumulating organisms comprising, in their genome, a PDAT gene transferred by recombinant DNA technology or somatic hybridization.

- 13. Transgenic oil accumulating organisms according to item 12 comprising, in their genome, a PDAT gene having specificity for substrates with a particular uncommon fatty acid and the gene for said uncommon fatty acid.
- 14. Transgenic organisms according to item 12 or 13 which are selected from the group consisting of fungi, plants and animals.
- 15. Transgenic organisms according to item 12 or 13 which are selected from the group of agricultural plants.
  - 16. Transgenic organisms according to item 12 or 13 which are selected from the group of agricultural plants and where said PDAT gene is expressed under the control of a storage organ specific promotor.
  - 17. Transgenic organisms according to item 12 or 13 which are selected from the group of agricultural plants and where said PDAT gene is expressed under the control of a seed promotor.
  - 18. Oils from organisms according to item 12 17.
  - 19. A method for altering acyl specificity of a PDAT by alteration of the nucleotide sequence of a naturally occurring encoding gene and as a consequence of this alternation creating a gene encoding for an enzyme with novel acyl specifity.
  - 20. A protein encoded by a DNA molecule according to item 1 or a functional fragment thereof.
  - 21. A protein of item 20 designated phospholipid:diacylglycerol acyltransferase.
- 25 22. A protein of item 21 which has a distinct acyl specificity.
  - 23. A protein of item 13 having the amino acid sequence as set forth in SEQ, ID NO. 2, 13, 14 or 15 (and the proteins encoded by the fullength or partial genes set forth in SEQ. ID. NO. 1, 3, 4, 5, 7, 9, 10, 11 or 12) or an amino acid sequence with at least 30 % homology to said amino acid sequence.
- 30 24. A protein of item 23 isolated from Saccharomyces cereviseae.

## General methods:

5

10

15

20

25

30

Yeast strains and plasmids. The wild type yeast strains used were either FY1679 (MATα his3-Δ200 leu2-Δ1 trp1-Δ6 ura3-52) or W303-1A (MATa ADE2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1) (7). The YNR008w::KanMX2 disruption strain FVKT004-04C(AL), which is congenic to FY1679, was obtained from the Euroscarf collection (8). A 2751 bp fragment containing the YNR008w gene with 583 bp of 5' and 183 bp of 3' flanking DNA was amplified W303-1A genomic DNA using Tag polymerase from with 5'-TCTCCATCTTCTGCAAAACCT-3' and 5'-CCTGTCAAAAACCTTCTCCTC-3' as primers. The resulting PCR product was purified by agarose gel electrophoresis and cloned into the EcoRV site of pBluescript (pbluescript-pdat). For complementation experiments, the cloned fragment was released from pBluescript by HindIII-Sacl digestion and then cloned between the HindIII and Sac sites of pFL39 (9), thus generating pUS1. For overexpression of the PDAT gene, a 2202 bp EcoRI fragment from the pBluscript plasmid which contains only 24 bp of 5' flanking DNA was cloned into the BamHI site of the GAL1-TPK2 expression vector pJN92 (12), thus generating pUS4.

Microsomal preparations. Microsomes from developing seeds of sunflower (Helianthus annuus), Ricinus communis and Crepis palaestina were prepared using the procedure of Stobart and Stymne (11). To obtain yeast microsomes, 1g of yeast cells (fresh weight) was re-suspended in 8 ml of ice-cold buffer (20 mM Tris-Cl, pH 7.9, 10 mM MgCl<sub>2</sub>, 1 mM EDTA, 5 % (v/v) glycerol, 1 mM DTT, 0.3 M ammonium sulfate) in a 12 ml glass tube. To this tube, 4 ml of glass beads (diameter 0.45-0.5 mm) were added, and the tube was then heavily shaken (3 x 60 s) in an MSK cell homogenizer (B. Braun Melsungen AG, Germany). The homogenized suspension was centrifuged at 20,000 x g for 15 min at 6°C and the resulting supernatant was again centrifuged at 100,000 x g for 2 h at 6°C. The 100,000 x g pellet was resuspended in 0.1 M potassium

10



phosphate (pH 7.2), and stored at -80°C. It is subsequently referred to as the crude yeast microsomal fraction.

Lipid substrates. Radio-labeled ricinoleic (12-hydroxy-9-octadecenoic) and vernolic (12,13-epoxy-9-octadecenoic) acids were synthesized enzymatically from [1-14C]oleic acid and [1-14C]linoleic acid, respectively, by incubation with microsomal preparations from seeds of Ricinus communis and Crepis palaestina, respectively (12). The synthesis of phosphatidylcholines (PC) or phosphatidylethanolamines (PE) with <sup>14</sup>C-labeled acyl groups in the sn-2 position was performed using either enzymatic (13), or synthetic (14) acylation of [14C]oleic, [14C]ricinoleic, or [14C]vernolic acid. Dioleoyl-PC that was labeled in the sn-1 position was synthesized from sn-1-[14C]oleoyl-lyso-PC and unlabeled oleic acid as described in (14). Sn-1-oleoyl-sn-2-[14C]ricinoleoyl-DAG was synthesized from PC by the action of phospholipase C type XI from B. Cereus (Sigma Chemical Co.) as described in (15). Monovernoloyl- and divernoleoyl-DAG were synthesized from TAG extracted from seeds of Euphorbia lagascae, using the TAG-lipase (Rizhopus arrhizus, Sigma Chemical Co.) as previously described (16). Monoricinoleoyl-TAG was synthesized according to the same method using TAG extracted from Castor bean.

20

25

30

15

Lipid analysis. Total lipid composition of yeast were determined from cells harvested from a 40 ml liquid culture, broken in a glass-bead shaker and extracted into chloroform as described by Bligh and Dyer (17), and then separated by thin layer chromatography in hexane/diethylether/acetic acid (80:20:1) using pre-coated silica gel 60 plates (Merck). The lipid areas were located by brief exposure to I<sub>2</sub> vapors and identified by means of appropriate standards. Polar lipids, sterol-esters and triacylglycerols, as well as the remaining minor lipid classes, referred to as other lipids, were excised from the plates. Fatty acid methylesters were prepared by heating the dry excised material at 85 °C for 60 min in 2% (v/v) sulfuric acid in dry methanol. The methyl esters were extracted with hexane and analyzed by GLC through a 50 m

x 0.32 mm CP-Wax58-CB fused-silica column (Chrompack), with methylheptadecanoic acid as an internal standard. The fatty acid content of each fraction was quantified and used to calculate the relative amount of each lipid class. In order to determine the total lipid content, 3 ml aliquots from yeast cultures were harvested by centrifugation and the resulting pellets were washed with distilled water and lyophilized. The weight of the dried cells was determined and the fatty acid content was quantified by GLC-analyses after conversion to methylesters as described above. The lipid content was then calculated as nmol fatty acid (FA) per mg dry weight yeast.

10

15

20

25

30

Enzyme assays. Aliquots of crude microsomal fractions (corresponding to 10 nmol of microsomal PC) from developing plant seeds or yeast cells were lyophilized over night. <sup>14</sup>C-Labeled substrate lipids dissolved in benzene were then added to the dried microsomes. The benzene was evaporated under a stream of N<sub>2</sub>, leaving the lipids in direct contact with the membranes, and 0.1 ml of 50 mM potassium phosphate (pH 7.2) was added. The suspension was thoroughly mixed and incubated at 30°C for the time period indicated, up to 90 min. Lipids were extracted from the reaction mixture using chloroform and separated by thin layer chromatography in hexane/diethylether/acetic acid (35:70:1.5) using silica gel 60 plates (Merck). The radioactive lipids were visualized and quantified on the plates by electronic autoradiography (Instant Imager, Packard, US).

<u>Yeast cultivation.</u> Yeast cells were grown at 28°C on a rotatory shaker in liquid YPD medium (1% yeast extract, 2% peptone, 2% glucose), synthetic medium (18) containing 2% (v/v) glycerol and 2% (v/v) ethanol, or minimal medium (19) containing 16 g/l of glycerol.

The instant invention is further characterized by the following examples which are not limiting:



Acyl-CoA-independent synthesis of TAG by oil seed microsomes. A large number of unusual fatty acids can be found in oil seeds (20). Many of these fatty acids, such as ricinoleic (21) and vernolic acids (22), are synthesized using phosphatidylcholin (PC) with oleoyl or linoleoyl groups esterified to the sn-2 position, respectively, as the immediate precursor. However, even though PC can be a substrate for unusual fatty acid synthesis and is the major membrane lipids in seeds, unusual fatty acids are rarely found in the membranes. Instead, they are mainly incorporated into the TAG. A mechanism for efficient and selective transfer of these unusual acyl groups from PC into TAG must therefore exist in oil seeds that accumulate such unusual fatty acids. This transfer reaction was biochemically characterized in seeds from castor bean (Ricinus communis) and Crepis palaestina, plants which accumulate high levels of ricinoleic and vernolic acid, respectively, and sunflower (Helianthus annuus), a plant which has only common fatty acids in its seed oil. Crude microsomal fractions from developing seeds were incubated with PC having <sup>14</sup>C-labeled oleoyl, ricinoleoyl or vernoloyl groups at the sn-2 position. After the incubation, lipids were extracted and analyzed by thin layer chromatography. We found that the amount of radioactivity that was incorporated into the neutral lipid fraction increased linearly over a period of 4 hours (data not shown). The distribution of [14C]acyl groups within the neutral lipid fraction was analyzed after 80 min (Fig. 1). Interestingly the amount and distribution of radioactivity between diffferent neutral lipids were strongly dependent both on the plant species and on the type of [14C]acyl chain. Thus, sunflower microsomes incorporated most of the label into DAG, regardless of the type of [14C]acvl group. In contrast, R. communis microsomes preferentially incorporated [14C]ricinoleoyl and [14C]vernoloyl groups into TAG, while [14C]oleyl groups mostly were found in DAG. C. palaestina microsomes, finally, incorporated only [14C]vernolyol groups into TAG, with [14C]ricinoleyl groups being found mostly as free fatty acids, and [14C]oleyl groups in DAG. This shows that the high in vivo levels of ricinoleic acid and vernolic acid in the TAG pool of R. communis

16

10

15

20

25

PCT/EP00/02701

and C. palaestina, respectively, can be explained by an efficient and selective transfer of the corresponding acyl groups from PC to TAG in these organisms.

The in-vitro synthesis of triacylglycerols in microsomal preparations of developing castor bean is summarized in table 1.

PDAT: a novel enzyme that catalyzes acyl-CoA independent synthesis of TAG. It was investigated if DAG could serve both as an acyl donor as well as an acyl acceptor in the reactions catalyzed by the oil seed microsomes. Therefore, unlabeled divernoloyl-DAG was incubated with either sn-1-oleoylsn-2-[14C]ricinoleoyl-DAG or sn-1-oleoyl-sn-2-[14C]ricinoleoyl-PC presence of R. communis microsomes. The synthesis of TAG molecules containing both [14C]ricinoleoyl and vernoloyl groups was 5 fold higher when [14C]ricinoleoyl-PC served as acyl donor as compared to [14C]ricinoleoyl-DAG (fig.1B). These data strongly suggests that PC is the immediate acyl donor and DAG the acyl acceptor in the acyl-CoA-independent formation of TAG by oil seed microsomes. Therefore, this reaction is catalyzed by a new enzyme which we call phospholipid: diacylglycerol acyltransferase (PDAT).

PDAT activity in yeast microsomes. Wild type yeast cells were cultivated under conditions where TAG synthesis is induced. Microsomal membranes were prepared from these cells and incubated with sn-2-[14C]-ricinoleovI-PC and DAG and the <sup>14</sup>C-labeled products formed were analyzed. The PC-derived [14C]ricinoleoyl groups within the neutral lipid fraction mainly were found in free fatty acids or TAG, and also that the amount of TAG synthesized was dependent on the amount of DAG that was added to the reaction (Fig.2). The in vitro synthesis of TAG containing both ricinoleoyl and vernoloyl groups, a TAG species not present in vivo, from exogenous added sn-2-114Clricinoleovl-PC and unlabelled vernoloyi-DAG (Fig. 2, lane 3) clearly demonstrates the existence of an acyl-CoA-independent synthesis of TAG involving PC and DAG as

10

15

20

25

substrates in yeast microsomal membranes. Consequently, TAG synthesis in yeast can be catalyzed by an enzyme similar to the PDAT found in plants.

# The PDAT encoding gene in yeast.

A gene in the yeast genome (YNR008w) is known, but nothing is known about the function of YNR008w, except that the gene is not essential for growth under normal circumstances. Microsomal membranes were prepared from the yeast strain FVKT004-04C(AL) (8) in which this gene with unknown function had been disrupted. PDAT activity in the microsomes were assayed using PC with radiolabelled fatty acids at the sn-2 position. The activity was found to be completely absent in the disruption strain (Fig. 2 lane 4). Significantly, the activity could be partially restored by the presence of YNR008w on the single 2 lane 5). Moreover, acyl groups of copy plasmid pUS1 (Fig. phosphatidylethanolamine (PE) were efficiently incorporated into TAG by microsomes from the wild type strain whereas no incorporation occured from this substrate in the mutant strain (data not shown). This shows that YNR008w encodes a yeast PDAT which catalyzes the transfer of an acyl group from the sn-2 position of phospholipids to DAG, thus forming TAG. It should be noted that no cholesterol esters were formed from radioactive PC even in incubations with added ergosterols, nor were the amount of radioactive free fatty acids formed from PC affected by disruption of the YNR008w gene (data not shown). This demonstrates that yeast PDAT do not have cholesterol ester synthesising or phospholipase activities.

Increased TAG content in yeast cells that overexpress PDAT. The effect of overexpressing the PDAT-encoding gene was studied by transforming a wild type yeast strain with the pUS4 plasmid in which the gene is expressed from the galactose-induced GAL1:TPK2 promoter. Cells containing the empty expression vector were used as a control. The cells were grown in synthetic glycerol-ethanol medium, and expression of the gene was induced after either 2 hours (early log phase) or 25 hours (stationary phase) by the addition of

10

15

20

25

galactose. The cells were then incubated for another 21 hours, after which they were harvested and assays were performed. We found that overexpression of PDAT had no significant effect on the growth rate as determined by the optical density. However, the total lipid content, measured as µmol fatty acids per mg yeast dry weight, was 47% (log phase) or 29% (stationary phase) higher in the PDAT overexpressing strain than in the control. Furthermore, the polar lipid and sterolester content was unaffected by overexpression of PDAT. Instead, the elevated lipid content in these cells is entirely due to an increased TAG content (Fig. 3A,B). Thus, the amount of TAG was increased by 2-fold in PDAT overexpressing early log phase cells and by 40% in stationary phase cells. It is interesting to note that a significant increase in the TAG content was achieved by overexpressing PDAT even under conditions (i.e. in stationary phase) where DAGAT is induced and thus contributes significantly to TAG synthesis. In vitro PDAT activity assayed in microsomes from the PDAT overexpressing strain was 7-fold higher than in the control strain, a finding which is consistent with the increased levels of TAG that we observed in vivo (Fig. 3C). These results clearly demonstrate the potential use of the PDAT gene in increasing the oil content in transgenic organisms.

Substrate specificity of yeast PDAT. The substrate specificity of yeast PDAT was anaiyzed using microsomes prepared from the PDAT overexpressing strain (see Fig. 4). The rate of TAG synthesis, under conditions given in figure 4 with di-oleoyl-PC as the acyl-donor, was 0.15 nmol per min and mg protein. With both oleoyl groups of PC labeled it was possible, under the given assay conditions, to detect the transfer of 11 pmol/min of [14C]oleovi chain into TAG and the formation of 15 pmol/min of lyso-PC. In microsomes from the PDAT-deficient strain, no TAG at all and only trace amounts of lyso-PC was detected, strongly suggesting that yeast PDAT catalyses the formation of equimolar amounts of TAG and lyso-PC when supplied with PC and DAG as substrates. The fact that somewhat more lyso-PC than TAG is formed can be

5

10

15

20

25

10

15

20

25

30



explained by the presence of a phospholipase in yeast microsomes, which produces lyso-PC and unesterified fatty acids from PC.

The specificity of yeast PDAT for different acyl group positions was investigated by incubating the microsomes with di-oleoyl-PC carrying a [14C]acyl group either at the sn-1 position (Fig. 4A bar 2) or the sn-2 position (Fig. 4A bar 3). We found that the major <sup>14</sup>C-labeled product formed in the former case was lyso-PC, and in the latter case TAG. We conclude that yeast PDAT has a specificity for the transfer of acyl groups from the sn-2 position of the phospholipid to DAG, thus forming sn-1-lyso-PC and TAG. Under the given assay conditions, trace amounts of 14C-labelled DAG is formed from the sn-1 labeled PC by the reversible action of a CDP-choline : choline phosphotransferase. This labeled DAG can then be further converted into TAG by the PDAT activity. It is therefore not possible to distinguish whether the minor amounts of labeled TAG that is formed in the presence of di-oleoyl-PC carrying a [14C]acyl group in the sn-1 position, is synthesized directly from the sn-1-labeled PC by a PDAT that also can act on the sn-1 postion, or if it is first converted to sn-1-labeled DAG and then acylated by a PDAT with strict selectivity for the transfer of acyl groups at the sn-2 position of PC. Taken together, this shows that the PDAT encoded by YNR008w catalyses an acyl transfer from the sn-2 position of PC to DAG, thus causing the formation of TAG and lyso-PC.

The substrate specificity of yeast PDAT was further analyzed with respect to the headgroup of the acyl donor, the acyl group transferred and the acyl chains of the acceptor DAG molecule. The two major membrane lipids of *S. cerevisiae* are PC and PE, and as shown in Fig. 4B (bars 1 and 2), dioleoyl-PE is nearly 4-fold more efficient than dioleoyl-PC as acyl donor in the PDAT-catalyzed reaction. Moreover, the rate of acyl transfer is strongly dependent on the type of acyl group that is transferred. Thus, a ricinoleoyl group at the *sn*-2 position of PC is 2.5 times more efficiently transferred into TAG than an oleoyl

group in the same position (Fig. 4B bars 1 and 3). In contrast, yeast PDAT has no preference for the transfer of vernoloyl groups over oleoyl groups (Fig. 4B bars 1 and 4). The acyl chain of the acceptor DAG molecule also affects the efficiency of the reaction. Thus, DAG with a ricinoleoyl or a vernoloyl group is a more efficient acyl acceptor than dioleoyl-DAG (Fig. 4B bars 1, 5 and 6). Taken together, these results clearly show that the efficiency of the PDAT-catalyzed acyl transfer is strongly dependent on the properties of the substrate lipids.

<u>PDAT genes.</u> Nucleotide and amino acid sequences of several PDAT genes are given as SEQ ID No. 1 through 15. Futher provisional and/or partial sequences are given as SEQ ID NO 1a through 5a and 1b through 11b, respectively. One of the Arabidopsis genomic sequences (SEQ ID NO. 4) identified an Arabidopsis EST cDNA clone; T04806. This cDNA clone was fully characterised and the nucleotide sequence is given as SEQ ID NO. 5. Based on the sequence homology of the T04806 cDNA and the *Arabidopsis thaliana* genomic DNA sequence (SEQ ID NO 4) it is apparent that an additional A is present at position 417 in the cDNA clone (data not shown). Excluding this nucleotide would give the amino acid sequence depicted in SEQ ID NO. 12.

Increased TAG content in seeds of Arabidopsis thaliana that express the yeast PDAT. For the expression of the yeast PDAT gene in Arabidopsis thaliana an EcoRI fragment from the pBluescript-PDAT was cloned together with napin promotor (25) into the vector pGPTV-KAN (26). A plasmid (pGNapPDAT) having the yeast PDAT gene in the correct orientation was identified and transformed into Agrobacterium tumefaciens. These bacteria were used to transform Arabidopsis thaliana columbia (C-24) plants using the root transformation method (27). Plants transformed with an empty vector were used as controls.

First generation seeds (T1) were harvested and germinated on kanamycin containing medium. Second generation seeds (T2) were pooled from individual plants and their fatty acid contents analysed by quantification of their methyl

5

10

15

20

25

5



esthers by gas liquid chromatography after methylation of the seeds with 2% sulphuric acid in methanol at 85 °C for 1,5 hours. Quantification was done with heptadecanoic acid methyl esters as internal standard.

From the transformation with pGNapPDAT one T1 plant (26-14) gave raise to seven T2 plants of which 3 plants yielded seeds with statistically (in a mean difference two-sided test) higher oil content than seeds from T2 plants generated from T1 plant 32-4 transformed with an empty vector (table 2).

10

20

30

## References cited in the description:

- 1. Bell, R. M. & Coleman, R. A. (1980) Annu. Rev. Biochem. 49, 459-487.
- 2. Stymne, S. & Stobart, K. (1987) in *The biochemistry of plants: a comprehensive treatsie, Vol. 9*, ed. Stumpf, P. K. (Academic Press, New York), pp. 175-214.
- 3. Cases, S. et al. (1998) Proc. Natl. Acad. Sci. U S A 95, 13018-13023.
- 4. Hobbs, D. H., Lu, C. & Hills, M. J. (1999) FEBS Lett. 452, 145-9
- 5. Zou, J., Wei, Y., Jako, C., Kumar, A., Selvaraj, G. & Taylor, D. C. (1999) Plant J. 19, 645-653.
- 6. Lardizabal, K., Hawkins, D., Mai, J., & Wagner, N. (1999) Abstract presented at the Biochem. Mol. Plant Fatty Acids Glycerolipids Symposium, South Lake Tahoe, USA.
- 7. Thomas, B. J. & Rothstein, R. (1989) Cell 56, 619-630.
- 15 8. Entian, K.-D. & Kötter, P. (1998) Meth. Microbiol. 26, 431-449.
  - 9. Kern, L., de Montigny, J., Jund, R. & Lacroute, F. (1990) Gene 88, 149-157.
  - 10. Ronne, H., Carlberg, M., Hu, G.-Z. & Nehlin, J. O. (1991) *Mol. Cell. Biol.* 11, 4876-4884.
  - 11. Stobart, K. & Stymne, S. (1990) in *Method in Plant Biochemistry, vol 4,* eds. Harwood, J. L. & Bowyer, J. R. (Academic press, London), pp. 19-46.
  - 12. Bafor, M., Smith, M. A., Jonsson, L., Stobrt, A. K. & Stymne, S. (1991) *Biochem. J.* **280**, 507-514.
  - 13. Banas, A., Johansson, I. & Stymne, S. (1992) Plant Science 84, 137-144.
  - 14. Kanda, P. & Wells, M. A. (1981) J. Lipid. Res. 22, 877-879.
  - 5 15. Ståhl, U., Ek, B. & Stymne, S. (1998) Plant Physiol. 117, 197-205.
    - 16. Stobart, K., Mancha, M. & Lenman M, Dahlqvist, A. & Stymne, S. (1997) *Planta* **203**, 58-66.
    - 17. Bligh, E. G. & Dyer, W. J. (1959) Can. J. Biochem. Physiol. 37, 911-917.
    - 18. Sherman, F., Fink, G. R. & Hicks, J. B. (1986) in *Laboratory Course Manual for Methods in Yeast Genentics* (Cold Spring Harbor Laboratory)
    - 19. Meesters, P. A. E. P., Huijberts, G. N. M. and Eggink, G. (1996) Appl. Microbiol. Biotechnol. 45, 575-579.
    - 20. van de Loo, F. J., Fox, B. G. & Sommerville, C. (1993), in *Lipid metabolism in plants*, ed. Moore, T. S. (CRC Press, Inc.), pp. 91-126.
- 21. van de Loo, F. J., Broun, P., Turner, S. & Sommerville, S. (1995) Proc. Natl.



- Acad. Sci. U S A 95, 6743-6747.
- 22. Lee, M., Lenman, M., Banas, A., Bafor, M., Singh, S., Schweizer, M., Nilsson, R., Liljenberg, C., Dahlqvist, A., Gummeson, P-O., Sjödahl, S., Green, A., and Stymne, S. (1998) *Science* **280**, 915-918.
- 23. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997) *Nucl. Acids Res.* **24**, 4876-4882.
  - 24. Saitou, N. & Nei, M. (1987) Mol. Biol. Evol. 4, 406-425.
  - 25. Stålberg, K., Ellerström, M., Josefsson, L., & Rask, L. (1993) *Plant Mol. Biol.* 23, 671
- 26. Becker, D., Kemper, E., Schell, J., Masterson, R. (1992) *Plant Mol. Biol.* 20, 1195
  - 27.D. Valvekens, M. Van Montagu, and Van Lusbettens (1988) Proc. Natl. Acad. Sci. U.S.A. 85, 5536

Description of Figures

### FIG. 1.

5

10

15

Metabolism of <sup>14</sup>C-labeled PC into the neutral lipid fraction by plant microsomes. (A) Microsomes from developing seeds of sunflower, R. communis and C. palaestina were incubated for 80 min at 30°C with PC (8 nmol) having oleic acid in its sn-1 position, and either <sup>14</sup>C-labeled oleic. ricinoleic or vernolic acid in its sn-2 position. Radioactivity incorporated in TAG (open bars), DAG (solid bars), and unsterified fatty acids (hatched bars) was layer chromatography followed by quantified using thin autoradiography, and is shown as percentage of added labeled substrate. (B) Synthesis in vitro of TAG carrying two vernoloyl and one [14C]ricinoleoyl group by microsomes from R. communis. The substrates added were unlabeled divernoloyl-DAG (5 nmol), together with either sn-1-oleoyl-sn-2-[14C]ricinoleovl-DAG (0.4 nmol, 7700 dpm/nmol) or sn-1-oleoyl-sn-2-[14C]ricinoleoyl-PC (0.4 nmol, 7700 dpm/nmol). The microsomes were incubated with the substrates for 30 min at 30°C, after which samples were removed for lipid analysis as described in the section "general methods". The data shown are the average of two experiments.

20

25

30

#### FIG. 2.

PDAT activity in yeast microsomes, as visualized by autoradiogram of neutral lipid products separated on TLC. Microsomal membranes (10 nmol of PC) from the wild type yeast strain FY1679 (lanes 1-3), a congenic yeast strain (FVKT004-04C(AL)) that is disrupted for YNR008w (lane 4) or the same disruption strain transformed with the plasmid pUS1, containing the YNR008w gene behind its native promotor (lane 5), were assayed for PDAT activity. As substrates, we used 2 nmol *sn*-1-oleoyl-*sn*-2-[<sup>14</sup>C]ricinoleoyl-PC together with either 5 nmol of dioleoyl-DAG (lanes 2, 4 and 5) or *rac*-oleoyl-vernoleoyl-DAG (lane 3). The enzymatic assay and lipid analysis was performed as described in Materials and Methods. The cells were precultured for 20 h in liquid YPD

WO 00/60095 26



medium, harvested and re-suspended in an equal volume of minimal medium (19) containing 16 g/l glycerol. The cells were then grown for an additional 24 h prior to being harvested. Selection for the plasmid was maintained by growing the transformed cells in synthetic medium lacking uracil (18). Abbreviations: 1-OH-TAG, monoricinoleoyl-TAG; 1-OH-1-ep-TAG, monoricinoleoyl-monovernoloyl-TAG; OH-FA, unesterified ricinoleic acid.

Fig. 3.

5

10

15

20

25

30

Lipid content (A.B) and PDAT activity (C) in PDAT overexpressing yeast cells. The PDAT gene in the plasmid pUS4 was overexpressed from the galactoseinduced GAL1-TPK2 promotor in the wild type strain W303-1A (7). Its expression was induced after (A) 2 hours or (B) 25 hours of growth by the addition of 2% final concentration (w/v) of galactose. The cells were then incubated for another 22 hours before being harvested. The amount of lipids of the harvested cells was determined by GLC-analysis of its fatty acid contents and is presented as µmol fatty acids per mg dry weight in either TAG (open bar), polar lipids (hatched bar), sterol esters (solid bar) and other lipids (striped bar). The data shown are the mean values of results with three independent yeast cultures. (C) In vitro synthesis of TAG by microsomes prepared from yeast cells containing either the empty vector (vector) or the PDAT plasmid (+ PDAT). The cells were grown as in Fig. 3A. The substrate lipids dioleoyl-DAG (2.5 nmol) and sn-1-oleoyl-sn-2-[14C]-oleoyl-PC (2 nmol) were added to aliquots of microsomes (10 nmol PC), which were then incubated for 10 min at 28 °C. The amount of label incorporated into TAG was quantified by electronic autoradiography. The results shown are the mean values of two experiments.

FIG. 4.

<u>Substrate specificity of yeast PDAT.</u> The PDAT activity was assayed by incubating aliquots of lyophilized microsomes (10 nmol PC) with substrate lipids at 30°C for 10 min (panel A) or 90 min (panel B). Unlabeled DAG (2.5 nmol) was used as substrates together with different labeled phospholipids, as shown



in the figure. (A) Sn-position specificity of yeast PDAT regarding the acyl donor substrate. Dioleoyl-DAG together with either sn-1-[14C]oleoyl-sn-2-[14C]oleoyl-PC (di-[14C]-PC), sn-1-[14C]oleovi-sn-2-oleovi-PC (sn1-[14C]-PC) or sn-1-oleovisn-2-[14C]oleoyl-PC (sn2-[14C]-PC). (B) Specificity of yeast PDAT regarding phospholipid headgroup and of the acyl composition of the phospholipid as well as of the diacylglycerol. Dioleoyl-DAG together with either sn-1-oleoyl-sn-2-[14C]oleoyl-PC (oleoyl-PC), sn-1-oleoyl-sn-2-[14C]oleoyl-PE (oleoyl-PE), sn-1-(ricinoleoyl-PC) oleoyl-sn-2-[14C]ricinoleoyl-PC or sn-1-oleoyl-sn-2-<sup>14</sup>Clvernoloyl-PC (vernoloyl-PC). In the experiments presented in the 2 bars to the far right, monoricinoleoyl-DAG (ricinoleoyl-DAG or mono-vernoloyl-DAG (vernolovI-DAG) were used together with sn-1-oleovI-sn-2-[14C]-oleovI-PC. The label that was incorporated into TAG (solid bars) and lyso-PC (LPC, open bars) was quantified by electronic autoradiography. The results shown are the mean values of two experiments. The microsomes used were from W303-1A cells overexpressing the PDAT gene from the GAL1-TPK2 promotor, as described in Fig. 3. The expression was induced at early stationary phase and the cells were harvested after an additional 24 h.

#### TAB.1:

5

15

20

25

30

In vitro synthesis of triacylglycerols in microsomal preparations of developing castor bean. Aliquots of microsomes (20 nmol PC) were lyophilised and substrate lipids were added in benzene solution: (A) 0.4 nmol [14C]-DAG (7760 dpm/nmol) and where indicated 1.6 nmol unlabelled DAG; (B) 0.4 nmol [14C]-DAG (7760 dpm/nmol) and 5 nmol unlabelled di-ricinoleoyl-PC and (C) 0.25 nmol [14C]-PC (4000 dpm/nmol) and 5 nmol unlabelled DAG. The benzene was evaporated by N<sub>2</sub> and 0.1 ml of 50 mM potassium phosphate was added, thoroughly mixed and incubated at 30 °C for (A) 20 min.; (B) and (C) 30 min.. Assays were terminated by extraction of the lipids in chloroform. The lipids were then separated by thin layer chromatography on silica gel 60 plates



(Merck; Darmstadt, Germany) in hexan/diethylether/acetic 35:70:1.5. The radioactive lipids were visualised and the radioactivity quantified on the plate by electronic autoradiography (Instant Imager, Packard, US). Results are presented as mean values of two experiments.

5

Radioactivity in different triacylglycerols (TAG) species formed. Abbreviations used: 1-OH-, mono-ricinoleoyl-; 2-OH, di-ricinoleoyl-; 3-OH-, triricinoleoyl; 1-OH-1-ver-, mono-ricinoleoly-monovernoleoyl-; 1-OH-2-ver-, mono-ricinoleoyl-divernoleoyl-. Radiolabelled DAG and PC were prepared enzymatically. The radiolabelled ricinoleoyl group is attached at the sn-2-position of the lipid and unlabelled oleoyl group at the sn-1-position. Unlabelled DAG with vernoleoyl- or ricinoleoyl chains were prepared by the action of TAG lipase (6) on oil of Euphorbia lagascae or Castor bean, respectively. Synthetic di-ricinoleoyl-PC was kindly provided from Metapontum Agribios (Italy).

15

20

10

## TAB.2:

Total fatty acids per mg of T2 seeds pooled from individual *Arabidopsis thaliana* plants transformed with yeast PDAT gene under the control of napin promotor (26-14) or transformed with empty vector (32-4).

\* = stastistical difference between control plants and PDAT transformed plants in a mean difference two-sided test at  $\alpha = 5$ .

15

20

## Description of the SEQ ID:

SEQ ID NO. 1: Genomic DNA sequence and suggested amino acid sequence of the Saccharomyces cerevisiae PDAT gene, YNR008w, with GenBank accession number Z71623 and Y13139, and with nucleotide ID number 1302481.

SEQ ID NO. 2: The amino acid sequence of the suggested open reading frame YNR008w from Saccharomyces cerevisiae.

SEQ ID NO. 3: Genomic DNA sequence of the Schizosaccharomyces pombe gene SPBC776.14.

SEQ ID NO. 4: Genomic DNA sequence of part of the Arabidopsis thaliana locus with GenBank accession number AB006704.

SEQ ID NO. 5: Nucleotide sequence of the Arabidopsis thaliana cDNA clone with GenBank accession number T04806, and nucleotide ID number 315966.

SEQ ID NO. 6: Predicted amino acid sequence of the Arabidopsis thaliana cDNA clone with GenBank accession number T04806.

SEQ ID NO. 7: Nucleotide and amino acid sequence of the Zea mays EST clone with GenBank accession number AI491339, and nucleotide ID number 4388167.

25 SEQ ID NO. 8: Predicted amino acid sequence of the Zea mays EST clone with GenBank accession number Al491339, and nucleotide ID number 4388167.

SEQ ID NO. 9: DNA sequence of part of the Neurospora crassa EST clone W07G1, with GenBank accession number Al398644, and nucleotide ID number 4241729.



SEQ ID NO. 10: Genomic DNA sequence of part of the Arabidopsis thaliana locus with GenBank accession number AC004557.

SEQ ID NO. 11: Genomic DNA sequence of part of the Arabidopsis thaliana locus with GenBank accession number AC003027.

SEQ ID NO. 12: DNA sequence of part of the Lycopersicon esculentum cDNA clone with GenBank accession number Al486635.

10 SEQ ID NO. 13: Amino acid sequence of the Schizosaccharomyces pombe putative open reading frame CAA22887 of the Schizosaccharomyces pombe gene SPBC776.14.

SEQ ID NO. 14: Amino acid sequence of the Arabidopsis thaliana putative open reading frame AAC80628 derived from the Arabidopsis thaliana locus with GenBank accession number AC004557.

SEQ ID NO 15: Amino acid sequence of the Arabidopsis thaliana putative open reading frame AAD10668 derived from the Arabidopsis thaliana locus with GenBank accession number AC003027.

Further provisional and/or partial sequences are defined through the following SEQ IDs:

25 SEQ ID NO. 1a: The amino acid sequence of the yeast ORF YNR008w from Saccharomyces cerevisiae.

SEQ ID NO. 2a: Amino acid sequence of the region of the Arabidopsis thaliana genomic sequence (AC004557).

15



SEQ ID NO. 3a: Amino acid sequence of the region of the Arabidopsis thaliana genomic sequence (AB006704).

SEQ ID NO. 4a: The corresponding genomic DNA sequence and amino acid sequence of the yeast ORF YNROO8w from Saccharomyces cerevisiae.

SEQ ID NO. 5a: The amino acid sequence of the yeast ORF YNROO8w from Saccharomyces cerevisiae derived form the corresponding genomic DNA sequence.

10

SEQ ID NO. 1b: Genomic DNA sequence of the Saccharomyces cerevisiae PDAT gene, YNR008w, genebank nucleotide ID number 1302481, and the suggested YNR008w amino acid sequence.

15

SEQ ID NO. 2b: The suggested amino acid sequence of the yeast gene YNR008w from Saccharomyces cerevisiae.

SEQ ID NO. 3b: Genomic DNA sequence of the Schizosaccharomyces pombe gene SPBC776.14.

SEQ ID NO. 4b: Genomic DNA sequence of part of the Arabidopsis thaliana locus with genebank accession number AB006704.

25 SEQ ID NO. 5b: Nucleotide sequence and the corresponding amino acid sequence of the Arabidopsis thaliana EST-clone with genebank accession number T04806, and ID number 315966.

SEQ ID NO. 6b: Nucleotide and amino acid sequence of the Zea mays cDNA clone with genebank ID number 4388167.



SEQ ID NO. 7b: Amino acid sequence of the Zea mays cDNA clone with genebank ID number 4388167.

SEQ ID NO. 8b: DNA sequence of part of the Neurospora crassa cDNA clone WO7G1, ID number 4241729.

SEQ ID NO. 9b: Genomic DNA sequence of part of the Arabidopsis thaliana locus with genebank accession number AC004557.

10 SEQ ID NO. 10b: Genomic DNA sequence of part of the Arabidopsis thaliana locus with genebank accession number AC003027.

SEQ ID NO. 11b: DNA sequence of part of the Lycopersicon esculentum cDNA clone with genebank accession number Al486635.

15

20

25



## **Claims**

5

10

20

- An enzyme catalysing in an acyl-CoA-independent reaction the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol.
- 2. An enzyme according to claim 1, comprising an amino acid sequence as set forth in SEQ ID No. 2 or a functional fragment, derivate, allele, homolog or isoenzyme thereof.
- 3. An enzyme according to claims 1 or 2 designated as phospholipid:diacylglycerol acyltransferase (PDAT).
- 4. An enzyme according to claims 1 to 3, comprising an amino acid sequence as set forth in SEQ ID No. 1a, 2b or 5a or a functional fragment, derivate, allele, homolog or isoenzyme thereof.
  - 5. An enzyme according to claims 1 to 4, comprising an amino acid sequence selected from the group consisting of sequences as set forth in SEQ ID No. 2a, 3a, 5b, 6, 7b, 8, 13, 14 or 15 or a functional fragment, derivate, allele, homolog or isoenzyme thereof.
  - 6. An enzyme according to claims 1 to 5, comprising an amino acid sequence encoded through a nucleotide sequence, a portion, derivate, allele or homolog thereof selected from the group consisting of sequences as set forth in SEQ ID No. 1, 1b, 3, 3b, 4, 4a, 4b, 5, 5b, 6b, 7, 8b, 9, 9b, 10, 10b, 11, 11b or 12 or a functional fragment, derivate, allele, homolog or isoenzyme of the enzyme encoding amino acid sequence.
- 7. A nucleotide sequence encoding an enzyme catalysing in an acyl-CoAindependent reaction the transfer of fatty acids from phospholipids to

10

15

20



diacylglycerol in the biosynthetic pathway for the production of triacylglycerol.

- 8. A nucleotide sequence according to claim 7 encoding an enzyme designated as phospholipid:diacylglycerol acyltransferase (PDAT).
  - 9. A nucleotide sequence according to claims 7 or 8, selected from the group consisting of sequences as set forth in SEQ ID No. 1, 1b, 3, 3b, 4, 4a, 4b, 9b, 10, 10b or 11 or a portion, derivate, allele or homolog thereof.

10. A partial nucleotide sequence corresponding to a fullength nucleotide sequence according to claims 7 to 9, selected from the group consisting of sequences as set forth in SEQ ID No. 5, 5b, 6b, 7, 8b, 9, 11b or 12 or a portion, derivate, allele or homolog thereof.

11. A nucleotide sequence according to claims 7 to 10, comprising a nucleotide sequence which is at least 40% homologous to a nucleotide sequence selected form the group consisting of those sequences set forth in SEQ ID No. 1, 1b, 3, 3b, 4, 4a, 4b, 5, 5b, 6b, 7, 8b, 9, 9b, 10, 10b, 11, 11b or 12.

- 12. A gene construct comprising a nucleotide sequence according to claims 7 to 11 operably linked to a heterologous nucleic acid.
- 13. A vector comprising a nucleotide sequence according to claims 7 to 11 or agene construct according to claim 12.
  - 14. A vector according to claim 13, which is an expression vector.
- 15. A vector according to claims 13 or 14, further comprising a selectable marker gene and/or nucleotide sequences for the replication in a host cell or the integration into the genome of the host cell.

5

20

- 16. A transgenic cell or organism containing a nucleotide sequence according to claims 7 to 11 and/or a gene construct according to claim 12 and/or a vector according to claims 13 to 15.
- 17. A transgenic cell or organism according to claim 16 which is an eucaryotic cell or organism.
- 18. A transgenic cell or organism according to claims 16 or 17 which is a yeast cell or a plant cell or a plant.
  - 19. A transgenic cell or organism according to claims 16 to 18 having an altered biosynthetic pathway for the production of triacylglycerol.
- 15 20. A transgenic cell or organism according to claims 16 to 19 having an altered oil content.
  - 21. A transgenic cell or organism according to claims 16 to 20 wherein the activity of PDAT is altered.
  - 22. A transgenic cell or organism according to claims 16 to 21 wherein the altered activity of PDAT is characterized by an alteration in gene expression, catalytic activity and/or regulation of activity of the enzyme.
- 23. A transgenic cell or organism according to claims 16 to 22 wherein the altered biosynthetic pathway for the production of triacylglycerol is characterized by the prevention of accumulation of undesirable fatty acids in the membrane lipids.
- 30 24. A process for the production of triacylglycerol, comprising growing a transgenic cell or organism according to claims 16 to 23 under conditions



whereby the said nucleotide sequence according to claims 7 to 11 is expressed.

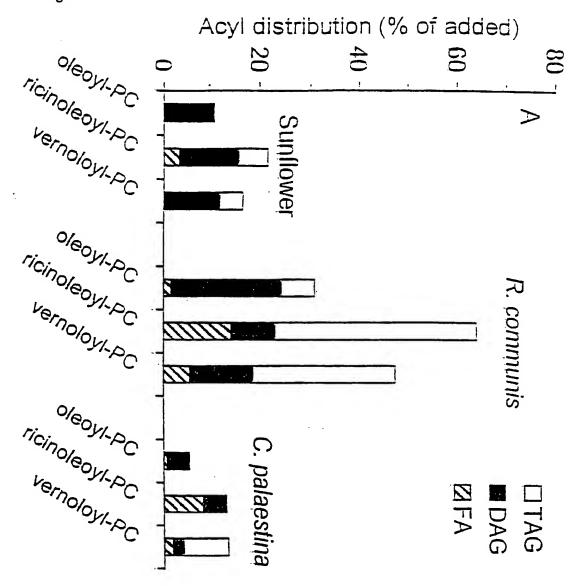
- 25. Triacylglycerols produced by a process according to claim 24.
- 26. Use of a nucleotide sequence according to claims 7 to 11 and/or an enzyme according to claims 1 to 6 for the production of triacylglycerol and/or triacylglycerols with uncommon fatty acids.
- 27. Use of a nucleotide sequence according to claims 7 to 11 and/or an enzyme according to claims 1 to 6 for the transformation of any cell or organism in order to be expressed in this cell or organism and result in an altered, preferably increased oil content of this cell or organism.

15

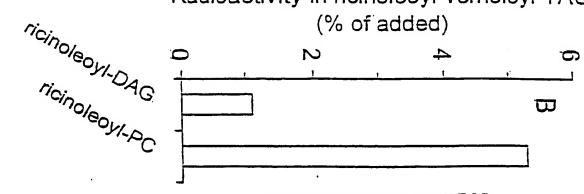
5

**Figurs** 

Fig. 1:

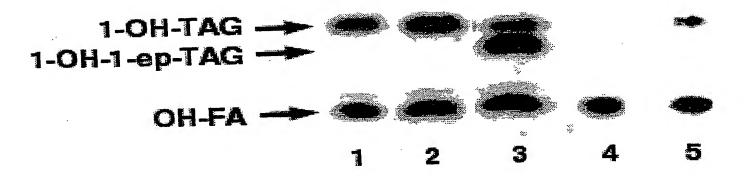


Radioactivity in ricinoleoyl-vernoloyl-TAG

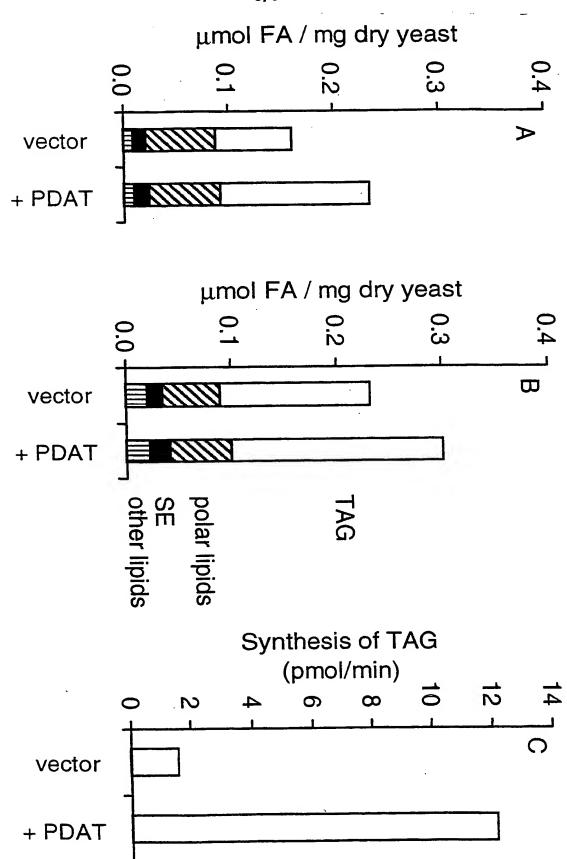




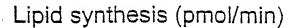
# Fig 2

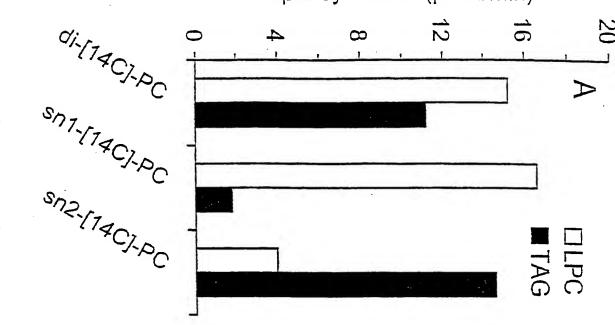




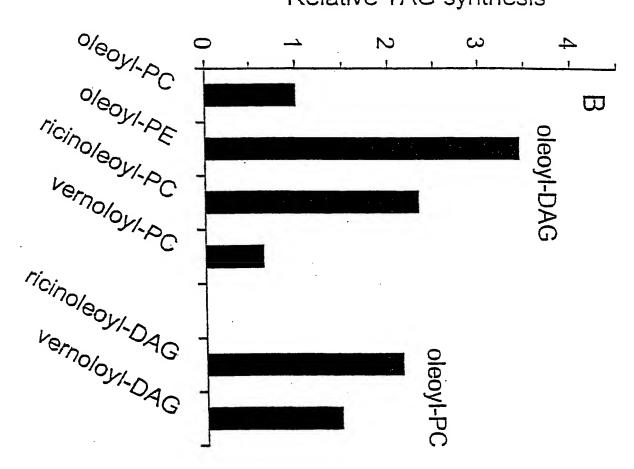








# Relative TAG-synthesis



					5/6					
3-OH-TAG	1	•	1		4,7	5,0	1,9	9,5	7,4	. 8'9
1-0H-2-ver-TAG	i	t	1,2	ı	•		•	ı	0,5	8,4
1-OH-1-ver-TAG	ı	1,3	0,5	•		•	•		6,01	<del>-</del> -
2-OH-TAG	12,4	12,1	10	24,8	8,0	8'6	16,7	9,4	11,5	10,8
1-OH-TAG	2,8	3,2	4	6,0	8'9	9,8	2,7	4,5	0,9	2'9
uniabelled lipid <sup>(2)</sup>	mono-ricinoleoyl-DAG	mono-vernoleoyl-DAG	di-vernoleoyl-DAG	di-ricinoleoyi-PC	none	di-oleoyl-DAG	mono-ricinoleoyl-DAG	di-ricinoleoyl-DAG	mono-vernoleoyi-DAG	di-vernoleoyl-DAG
Substrate added [14C]-lipid <sup>(2)</sup>	A mono-i 14C]-ricinoleoyl-DAG	A mono-[14C]-ricinoleoyl-DAG	A mono-[14C]-ricinoleoyi-DAG	A mono-[ <sup>14</sup> C]-ricinoleoyl-DAG	B mono-[ <sup>14</sup> C]-ricinoleoyi-PC	C mono-f 14Cj-ricinoleoyl-PC	C mono-f 14C]-ricinoleoyl-PC	C mono-f 14C1-ricinoleovI-PC	C mono-[ <sup>14</sup> C]-ricinoleoyl-PC	C mono-[14C]-ricinoleoyl-PC
	unlabelled lipid <sup>(2)</sup> 1-OH-TAG 2-OH-TAG 1-OH-1-ver-TAG 1-OH-2-ver-TAG	1-OH-TAG 2-OH-TAG 1-OH-1-ver-TAG 1-OH-2-ver-TAG 2,8 12,4 -	1-OH-TAG 2-OH-TAG 1-OH-1-ver-TAG 1-OH-2-ver-TAG 2,8 12,4	unlabelied lipid <sup>(2)</sup> 1-OH-TAG         2-OH-TAG         1-OH-1-ver-TAG         1-OH-2-ver-TAG           a mono-ricinoleoyl-DAG         2,8         12,4         -         -           a mono-vernoleoyl-DAG         3,2         12,1         1,3         -           a di-vernoleoyl-DAG         4         10         0,5         1,2	1-OH-TAG       2-OH-TAG       1-OH-1-ver-TAG       1-OH-2-ver-TAG         2,8       12,4       -       -         3,2       12,1       1,3       -         4       10       0,5       1,2         0,3       24,8       -       -	1-OH-TAG       2-OH-TAG       1-OH-1-ver-TAG       1-OH-2-ver-TAG       3-OH-TAG         2,8       12,4       -       -       -         3,2       12,1       1,3       -       -         4       10       0,5       1,2       -         0,3       24,8       -       -       -         6,8       8,0       -       -       4,7	1-OH-TAG       2-OH-TAG       1-OH-1-ver-TAG       1-OH-2-ver-TAG       3-OH-TAG         2,8       12,4       -       -       -         3,2       12,1       1,3       -       -         4       10       0,5       1,2       -         0,3       24,8       -       -       -         6,8       8,0       -       -       4,7         8,6       9,8       -       5,0	2.0H-TAG       2-OH-TAG       1-OH-1-ver-TAG       1-OH-2-ver-TAG       3-OH-TAG         2,8       12,4       -       -       -         3,2       12,1       1,3       -       -         4       10       0,5       1,2       -       -         0,3       24,8       -       -       -       -         6,8       8,0       -       -       4,7         8,6       9,8       -       -       5,0         5,7       16,7       -       -       1,9	1-OH-TAG       1-OH-TAG       1-OH-1-ver-TAG       1-OH-2-ver-TAG       3-OH-TAG         2,8       12,4       -       -       -         3,2       12,1       1,3       -       -         4       10       0,5       1,2       -       -         6,8       8,0       -       -       4,7         8,6       9,8       -       -       4,7         5,7       16,7       -       1,9         4,5       9,4       -       -       9,5	1-OH-TAG         2-OH-TAG         1-OH-i-ver-TAG         1-OH-2-ver-TAG         3-OH-TAG           2,8         12,4         -         -         -           3,2         12,1         1,3         -         -           4         10         0,5         1,2         -           0,3         24,8         -         -         -           6,8         8,0         -         -         4,7           8,6         9,8         -         -         1,9           5,7         16,7         -         -         1,9           4,5         9,4         -         -         9,5           6,0         11,5         10,9         0,5         7,4



Tab. 2:

T1 plant deviation	T2 plant number	nmol fatty acids per mg seed	standard
32-4	1	1277	±11 (n=2)
	4	1261	±63 (n=3)
	5	1369	$\pm 17 \text{ (n=3)}$
	6	1312	±53 (n=4)
	7	1197	±54 (n=5)
	8	1240	<u>+</u> 78 (n=4)
	9	1283	$\pm 54 \ (n=5)$
	10	1381	<u>+</u> 35 (n=5)
26-14	1	1444	±110 (n=4)
20 1.	2	1617*	$\pm 109 (n=4)$
	3	1374	$\pm 37 (n=2)$
	5	1562*	±70 (n=4)
	6	1393	$\pm$ 77 (n=4)
	7	1433	<u>+</u> 98 (n=4)
	8	1581*	±82 (n=4)



# Sequence Listing

	> 198 > ge	nomi	c DN		cere	visi	ae									
<221 <222			1983	)												
	ggc										caa Gln					48
											aag Lys					96
											aag Lys					144
											aaa Lys . 60					192
						Arg					gat Asp					240
										Leu	ttg Leu				Phe	288
				Val					Ser		ttg Leu			Asn		336
			asp					Tyr			gat Asp		Lys		gtt Val	384
		Glr					Phe					Glr			aac Asn	432



						_	~		•	_			_	gtt Val		480
				_						_				gtt Val 175	_	528
_	_			- ·			•			-	•			gtt Val		576
	-	_		_		_					_			ctg Leu		624
														tgt Cys		672
_														ccg Pro		720
										Ser				ttc Phe 255	Ile	768
_				Ile					Phe					gta Val		816
			Pro					Ser					Trp	agg Arg		864
		Leu					Arg					Thr		rcta Leu	aag Lys	912
	Gln					His					Glu			tgt Cys	tta Leu 320	960
					Gly					Phe				aaa Lys		1008





								3 / 5	53							
				ggc Gly												1056
			340					345					350			
		His		gat Asp			Ile					Thr				1104
act	663	355	aca	gtt	CC3	act	360	255	agt	aat	αaa	365 ata	888	gat	acc	1152
				Val												
				acg Thr		_	_									1200
	att	gag	r aga	gta		atg	tta	caa	acg		ggt	ggt	ata	cca		1248
Arg	Ile	Glu	a Arg	Val 405	_	Met	Leu	Gln	Thr 410		Gly	Gly	Ile	Pro 415		
			Lys						Trp					Ser	tct Ser	1296
tca	. gac	r gat	420 c gca		r aat	aac	aac			aca	a tac	ggc			att	1344
			o Ala					Thr					Asn		e Ile	
_		_		_	_	_	_	_							a atg	1392
	450					455					460	-				
	s As					Thi					r Pro				c caa 1 Gln 480	1440
					u Glr					у Ту					a gaa u Glu 5	
па	7 FF	a an	a aa			t ct:	a ca	c ca			c ta	a tc	o aa		a atg	1536
				s As					s Ly					n Pr	o Met	
			o Le					o Hi					r Cy		a tac e Tyr	

WO 00/60095 4 / 53

660

PCT/EP00/02701

								4 /	<i>33</i>						
													gag Glu		1632
-			_										caa Gln		1680
													cat His 575		1728
							_	_					gcc Ala	_	1776
		_				-	_						ttt Phe	-	1824
	_	Gly		-		_	Ala	_		_	_	Ile		agc Ser	1872
	Glu	_		_		Ile	_			_	Ser			gat Asp 640	1920
					Glr					ser				cag Gln	1968
_	ccc Pro			_		<b>a</b>									1986



<210> 2

<211> 661

<212> PRT

<213> Saccharomyces cerevisiae

165

<400>2

Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg

20 25

Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Gly 40

Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg

Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu 70 75

Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe 85

Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe 110 100 105

Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val 120 125

Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn 135

Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly 145 150 155

Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val

170

Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile 185

Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp 200

Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp 215

Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn 235 230

Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile 245 250

Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile 265

Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu

280

Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys 295

Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu 310 315

Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp 325 330 335



Val	Glu	Ala	Glu	Gly	Pro	Leu	Tyr	Gly	Asn	Gly	Gly	Arg	Gly	Trp	Val
			340					345					350		
Asn	Glu	His	Ile	Asp	Ser	Phe	Ile	Asn	Ala	Ala	Gly	Thr	Leu	Leu	Gly
		355					360					365			
Ala	Pro	Lys	Ala	Val	Pro	Ala	Leu	Ile	Ser	Gly	Glu	Met	Lys	Asp	Thr
	370					375					380				
Ile	Gln	Leu	Asn	Thr	Leu	Ala	Met	Tyr	Gly	Leu	Glu	Lys	Phe	Phe	Ser
385					390					395					400
Arg	Ile	Glu	Arg	Val	Lys	Met	Leu	Gln	Thr	Trp	Gly	Gly	Ile	Pro	Ser
				405					410					415	
Met	Leu	Pro	Lys	Gly	Glu	Glu	Val	Ile	Trp	Gly	Asp	Met	Lys	Ser	Ser
			420					425		_	-		430		
Ser	Glu	Asp	Ala	Leu	Asn	Asn	Asn	Thr	Asp	Thr	Tyr	Gly	Asn	Phe	Ile
		435					440		-		-	445			
Arg	Phe	Glu	Arg	Asn	Thr	Ser	Asp	Ala	Phe	Asn	Lys	Asn	Leu	Thr	Met
_	450		_			455	_				460				
Lys		Ala	Ile	Asn	Met	Thr	Leu	Ser	Ile	Ser	Pro	Glu	Trp	Leu	Gln
465	_				470					475			•		480
Arg	Arg	Val	His	Glu		Tvr	Ser	Phe	Glv		Ser	Lvs	Asn	Glu	
•				485					490	_		-2 -		495	
Glu	Leu	Ara	Lvs		Glu	Leu	His	His			Trp	Ser	Asn	-	Met
		-	500					505					510		
Glu	Val	Pro			Glu	Ala	Pro			Lvs	Ile	Tvr		Ile	Tvr
		515					520			-3 -		525	-		-1-
Glv	Val			Pro	Thr	Glu	Arg		Tvr	Val	Tvr			Glu	Asp
-	530					535			•		540				
Asp	Ser	Ser	Ala	Leu	Asn			Ile	Asp	Tvr		Ser	· Lvs	Gln	Pro
545					550					555			-1-		560
		Leu	Thr	Glu			Gly	Thr	Val			Va 1	Ala	His	
				565			4		570					575	
Met	Cvs	His	Lvs			Gln	Glv	Ala			Tvr	Asn	Pro		Gly
	-		580					585			-1-		590		<b>U</b> -1
Ile	Asn	. Val			Val	Glu	Met			Glr	Pro	Ast		Phe	Asp
		595				-	600					605			1100
Ile	Aro			r Ala	Lvs	Ser			His	: Val	Asr			Glv	Ser
	610		0-1			615				, , , , ,	620				
Ala			ιλετ	. Acr	. Тч <i>гэ</i> г			Taze	T16	. <b>λ</b> 1ε			r Aen	Gly	Asp
625				101	630		. <u></u>	. Lys		635		JLY	HOL	. Сту	640
		Gla	Pro	λ χ			507	. Ac+	1 T.A1			ጥም	\ \T=1	S=~	Gln
				645		. Dec	. Set	. ASI	650		. 311.		, A CT T	655	
Mot	. Dr.	Phe	D~~	_					0.50	•				455	
		, E116	660		•										
			900	,											







~?

<210> 3

<211> 2312

<212> genomic DNA

<213> Schizosaccharomyces pombe

-	_	_	•	-	•
A	Т	G	G	CG	;
3	m	$\overline{}$	m	~	

<400> 3					
ATGGCGTCTT CCAAGAAGAG	CAAAACTCAT	******	A A C A A C T C A A		50
ATCTCCTATC GACTTACCAA					100
AGCAACCTTC AGCGTCCGAA					150
TCTAAATTTG GAAAAAGATT					200
ATGCGGTGCT TTTTTTTCG					250
CTGCTACGTT AGATAAATTT					300
GATGACATTA AAGGATATTT					350
TACTACGGAC AAGCCTTCGC					400
GTCTTGATAT GTACAATGAG					450
GTTCCTGGTG TTATCAGCTC					500
CTCGATTCCT TACTTTAGGA					550
AGGCAATGTT CCTTGACAAG					600
AAAAAACCG GCTTGGATCC					650
GTTTGAAGCA GCTGATTTTT					700
TAATTGAAAA CCTTGCTGCA					750
GCTTCTTACG ATTGGCGGTT					800
ATATTTTCA AAGTTAAAAA					850
AGAAAAAGGT AGTGTTGATT					900
TATTTTTTA AGTGGGTTGA					950
TTGGGTTAAT GATCATATTG					1000
GTTTGACTAC GTTTCTAACT				1050	
GCACCCAAAA CAGTGGCAGC	GCTTTTATCG	GGTGAAATGA	AAGATACAGG		1100
TATTGTAATT ACATTAAACA	TGTTAATATT	TAATTTTTGC	TAACCGTTTT		1150
AAGCTCAATT GAATCAGTTT					1200
GATTTGTTAC TAATTTACTG	TTTAGTTTGG	AAAAATTTTT	TTCCCGTTCT		1250
GAGGTATATT CAAAAATACA	. AATGTGCTCT	ACTTTTTCTA	ACTTTTAATA		1300
GAGAGCCATG ATGGTTCGCA	. CTATGGGAGG	AGTTAGTTCT	ATGCTTCCTA		1350
AAGGAGGCGA TGTTGTATGG	GGAAATGCCA	GTTGGGTAAG	AAATATGTGC		1400
TGTTAATTTT TTATTAATAT	TTAGGCTCCA	GATGATCTTA	ATCAAACAAA		1450
TTTTTCCAAT GGTGCAATTA	TTCGATATAG	AGAAGACATT	GATAAGGACC		1500
ACGATGAATT TGACATAGAT	GATGCATTAC	AATTTTTAAA	AAATGTTACA		1550
GATGACGATT TTAAAGTCAT	GCTAGCGAAA	AATTATTCCC	ACGGTCTTGC		1600
TTGGACTGAA AAAGAAGTGT	AATAAAAAT C	CGAAATGCCG	TCTAAATGGA	1650	
TAAATCCGCT AGAAGTAAGA	ACATTAAAGT	TACTAAATTA	TACTAACCCA		1700
AATAGACTAG TCTTCCTTAT	GCTCCTGATA	TGAAAATTTA	TTGCGTTCAC		1750
GGGGTCGGAA AACCAACTGA	GAGAGGTTAT	TATTATACTA	ATAATCCTGA		1800
GGGGCAACCT GTCATTGATT	CCTCGGTTAA	TGATGGAACA	AAAGTTGAAA		1850
ATGTGAGAGA ATTTATGTTT	CAAACATTCT	ATTAACTGTT	TTATTAGGGT		1900
ATTGTTATGG ATGATGGTGA	TGGAACTTTA	CCAATATTAG	CCCTTGGTTT		1950
GGTGTGCAAT AAAGTTTGG	C AAACAAAAAG	GTTTAATCCT	GCTAATACAA		2000
GTATCACAAA TTATGAAATO					2050
GGAGGACCTC GCTCGGCAGA					2100
AAATGTATGT TCATTTTACO					2150
AAGGAAATTA TTTTAAAAG				2200	
CCGTTATATA TCAGATATCO					2250
TAACTAACTA ACCGAACAGO					2300
CCTAGAAATT AA		23			



2850

<210> 4 <211> 3685 <212> genomic DNA <213> Arabidopsis thaliana

ATGCCCCTTA TTCATCGGAA AAAGCCGACG GAGAAACCAT CGACGCCGCC 50 ATCTGAAGAG GTGGTGCACG ATGAGGATTC GCAAAAGAAA CCACACGAAT 100 CTTCCAAATC CCACCATAAG AAATCGAACG GAGGAGGGAA GTGGTCGTGC 150 ATCGATTCTT GTTGTTGGTT CATTGGGTGT GTGTGTGAA CCTGGTGGTT ATCGATTCTT GTTGTTGGTT CATTGGGTGT GTGTGTAA CCTGGTGGTT
TCTTCTCTCT CTTTACAACG CAATGCCTGC GAGCTTCCCT CAGTATGTAA
CGGAGCGAAT CACGGGTCCT TTGCCTGACC CGCCCGGTGT TAAGCTCAAA
AAAGAAGGTC TTAAGGCGAA ACATCCTGTT GTCTTCATTC CTGGGATTGT
CACCGGTGGG CTCGAGCTTT GGGAAGGCAA ACAATGCGCT GATGGTTTAT
TTAGAAAACG TTTGTGGGGT GGAACTTTTG GTGAAGTCTA CAAAAGGTGA
GCTCAACAAT TCTCACTCTT CCTTTATATT GGGATTTGGA TTGGATCTGA
TGTTTGTCTG TCTTACTCTT TACTTTTTT TTTTTTTGAT GTGAAACGCT
ATTTTCTTAA GAGACTATTT CTGTATGTGT AAGGTAAGCG TTCCAAGGAC
GTAATTGGCT TGGACTATTT CTGTTTGATT GTTAACTTTA GGATATAAAA
TAGCTGCCTT GGAATTTCAA GTCATCTTAT TGCCAAATCT GTTGCTAGAC
ATGCCCTAGA GTCCGTTCAT AACAAGTTAC TTCCTTTACT GTCGTTGCGT
GTAGATTTAG CTTTGTGTAG CGTATAATGA AGTAGTGTTT TATGTTTTGT 250 350 400 450 500 550 600 650 700 750 800 GTAGATTTAG CTTTGTGTAG CGTATAATGA AGTAGTGTTT TATGTTTTGT 850 TGGGAATAGA GAAGTTCTAA CTACATCTGT GGAAAGTGTG TTCAGGCTGT GATAGAGGAC TGTTGCTTTA TTATTCAACT ATGTATATGT GTAATTAAAG CTAGTTCCTT TTTGATCTTT CAGCTCAATG TGCTTTTCTC AATTTTTTC TCAATTTCAA AGTTTCACAT CGAGTTTATT CACATGTCTT GAATTTCGTC 900 950 1000 1050 CATCCTCGTT CTGTTATCCA GCTTTGAACT CCTCCCGACC CTGCTATGGA 1100 TATATTAAAA AAAAAGTGTT TTGTGGGTTG CATCTTTGTT ACGATCTGCA 1150 TCTTCTTCTT TCGGCTCAGT GTTCATGTTT TTGCTATGGT AGAGATGGGC 1200 AATGTTATTG TTGATGGTAA CAGTGGTATA GTTGATAGTA TCTTAACTAA
TCAATTATCT CTTTGATTCA GGCCTCTATG TTGGGTGGAA CACATGTCAC
TTGACAATGA AACTGGGTTG GATCCAGCTG GTATTAGAGT TCGAGCTGTA
1350 1250 1300 TCAGGACTCG TGGCTGCTGA CTACTTTGCT CCTGGCTACT TTGTCTGGGC 1400 AGTGCTGATT GCTAACCTTG CACATATTGG ATATGAAGAG AAAAATATGT 1450 ACATGGCTGC ATATGACTGG CGGCTTTCGT TTCAGAACAC AGAGGTTCTT 1500 TTCTCATCGT TCTTTCTATT ATTCTGTTCC ATGTTACGTT TCTTTCTTCA 1550 TTACTTAAGG CTTAAATATG TTTCATGTTG AATTAATAGG TACGTGATCA 1600 GACTCTTAGC CGTATGAAAA GTAATATAGA GTTGATGGTT TCTACCAACG 1650 GTGGAAAAAA AGCAGTTATA GTTCCGCATT CCATGGGGGT CTTGTATTTT CTACATTTTA TGAAGTGGGT TGAGGCACCA GCTCCTCTGG GTGGCGGGGG TGGGCCAGAT TGGTGTGCAA AGTATATTAA GGCGGTGATG AACATTGGTG GACCATTTCT TGGTGTTCCA AAAGCTGTTG CAGGGCTTTT CTCTGCTGAA GCAAAGGATG TTGCAGTTGC CAGGTATTGA ATATCTGCTT ATACTTTTGA 1900 TGATCAGAAC CTTGGCTCTG GAACTCAAAG TTATTCTACT AAATATCAAT 1950 TCTAATAACA TTGCTATATT ATCGCTGCAA CTGACATTGG TTGATTATTT 2000 TTGCTGCTTA TGTAACTGAA ACTCTCTTGA GATTAGACAA ATGATGAATT 2050 GATAATTCTT ACGCATTGCT CTGTGATGAC CAGTTTCTTA GCTTCGACGA 2100 TAACATTTGT CATACTGTCT TTTGGAGGGC ATTGAATTTT GCTATGGAAA 2150 GCGCTGGAGC TTCCATGCTT GCATTCTTTA CCAATTAGCG TTATTCTGCT 2200 TCTTTCAATT TTCTTGTATA TGCATCTATG GTCTTTTATT TCTTCTTAAT 2250 TAAAGACTCG TTGGATTAGT TGCTCTATTA GTCACTTGGT TCCTTAATAT 2300 AGAACTTTAC TTTCTTCGAA AATTGCAGAG CGATTGCCCC AGGATTCTTA 2350 GACACCGATA TATTTAGACT TCAGACCTTG CAGCATGTAA TGAGAATGAC ACGCACATGG GACTCAACAA TGTCTATGTT ACCGAAGGGA GGTGACACGA 2450 TATGGGGCGG GCTTGATTGG TCACCGGAGA AAGGCCACAC CTGTTGTGGG 2500 AAAAAGCAAA AGAACAACGA AACTTGTGGT GAAGCAGGTG AAAACGGAGT 2550 TTCCAAGAAA AGTCCTGTTA ACTATGGAAG GATGATATCT TTTGGGAAAG 2600 AAGTAGCAGA GGCTGCGCCA TCTGAGATTA ATAATATTGA TTTTCGAGTA 2650 AGGACATATA AATCATAATA AACCTTGTAC ATTTTGTGAT TGTATGATGA 2700 ATATCTGTAC ATTTTATCTG GTGAAGGGTG CTGTCAAAGG TCAGAGTATC 2750 CCAAATCACA CCTGTCGTGA CGTGTGGACA GAGTACCATG ACATGGGAAT 2800 TGCTGGGATC AAAGCTATCG CTGAGTATAA GGTCTACACT GCTGGTGAAG

WO 00/60095

9 / 53 PCT/EP00/02701

CTATAGATCT	ACTACATTAT	GTTGCTCCTA	AGATGATGGC	GCGTGGTGCC		2900
GCTCATTTCT	CTTATGGAAT	TGCTGATGAT	TTGGATGACA	CCAAGTATCA		2950
AGATCCCAAA	TACTGGTCAA	ATCCGTTAGA	GACAAAGTAA	GTGATTTCTT		3000
GATTCCAACT	GTATCCTTCG	TCCTGATGCA	TTATCAGTCT	TTTTGTTTTC		3050
GGTCTTGTTG	GATATGGTTT	TCAGCTCAAA	GCTTACAAAG	CTGTTTCTGA		3100
GCCTTTCTCA	AAAAGGCTTG	CTCAGTAATA	TTGAGGTGCT	AAAGTTGATA		3150
CATGTGACTC	TTGCTTATAA	ATCCTCCGTT	TGGTTTGTTC	TGCTTTTTCA		3200
GATTACCGAA	TGCTCCTGAG	ATGGAAATCT	ACTCATTATA	CGGAGTGGGG		3250
ATACCAACGG	AACGAGCATA	CGTATACAAG	CTTAACCAGT	CTCCCGACAG	3300	
TTGCATCCCC	TTTCAGATAT	TCACTTCTGC	TCACGAGGAG	GACGAAGATA		3350
GCTGTCTGAA	AGCAGGAGTT	TACAATGTGG	ATGGGGATGA	AACAGTACCC		3400
GTCCTAAGTG	CCGGGTACAT	GTGTGCAAAA	GCGTGGCGTG	GCAAGACAAG		3450
ATTCAACCCT	TCCGGAATCA	AGACTTATAT	AAGAGAATAC	AATCACTCTC		3500
CGCCGGCTAA	CCTGTTGGAA	GGGCGCGGA	CGCAGAGTGG	TGCCCATGTT		3550
GATATCATGG	GAAACTTTGC	TTTGATCGAA	GATATCATGA	GGGTTGCCGC		3600
CGGAGGTAAC	GGGTCTGATA	TAGGACATGA	CCAGGTCCAC	TCTGGCATAT		3650
TTGAATGGTC	GGAGCGTATT	GACCTGAAGC	TGTGA	3685		

10 / 53



<210> 5 <211> 2427 <212> cDNA

<213> Arabidopsis thaliana

<400> 5						
AGAAACAGCT	CTTTGTCTCT	CTCGACTGAT	CTAACAATCC	CTAATCTGTG		50
TTCTAAATTC		TTTGACAAAG	TCCGTATAGC	TTAACCTGGT		100
TTAATTTCAA		TGCCCCTTAT	TCATCGGAAA	AAGCCGACGG		150
AGAAACCATC	GACGCCGCCA		TGGTGCACGA			200
CAAAAGAAAC	CACACGAATC	TTCCAAATCC	CACCATAAGA			250
AGGAGGGAAG	TGGTCGTGCA			ATTGGGTGTG		300
TGTGTGTAAC	CTGGTGGTTT	CTTCTCTTCC	<del>-</del>	AATGCCTGCG		350
AGCTTCCCTC	AGTATGTAAC			TGCCTGACCC		400
GCCCGGTGTT	AAGCTCAAAA	AAAGAAGGTC	TTAAGGCGAA	ACATCCTGTT	450	
GTCTTCATTC	CTGGGATTGT	CACCGGTGGG	CTCGAGCTTT	GGGAAGGCAA		500
ACAATGCGCT	GATGGTTTAT	TTAGAAAACG	TTTGTGGGGT	GGAACTTTTG		550
GTGAAGTCTA	CAAAAGGCCT	CTATGTTGGG	TGGAACACAT	GTCACTTGAC		600
AATGAAACTG	GGTTGGATCC	AGCTGGTATT	AGAGTTCGAG	CTGTATCAGG		650
ACTCGTGGCT	GCTGACTACT	TTGCTCCTGG	CTACTTTGTC	TGGGCAGTGC		700
TGATTGCTAA	CCTTGCACAT	ATTGGATATG	AAGAGAAAAA	TATGTACATG		750
GCTGCATATG	ACTGGCGGCT	TTCGTTTCAG	AACACAGAGG	TACGTGATCA		800
GACTCTTAGC	CGTATGAAAA	GTAATATAGA	GTTGATGGTT	TCTACCAACG		850
GTGGAAAAAA	AGCAGTTATA	GTTCCGCATT	CCATGGGGGT	CTTGTATTTT		900
CTACATTTTA	TGAAGTGGGT	TGAGGCACCA	GCTCCTCTGG	GTGGCGGGG		950
TGGGCCAGAT	TGGTGTGCAA	AGTATATTAA	GGCGGTGATG	AACATTGGTG	1000	
GACCATTTCT	TGGTGTTCCA	AAAGCTGTTG	CAGGGCTTTT	CTCTGCTGAA		1050
GCAAAGGATG	TTGCAGTTGC	CAGAGCGATT	GCCCCAGGAT	TCTTAGACAC		1100
CGATATATTT	AGACTTCAGA	CCTTGCAGCA	TGTAATGAGA	ATGACACGCA		1150
CATGGGACTC	AACAATGTCT	ATGTTACCGA	AGGGAGGTGA	CACGATATGG		1200
GGCGGGCTTG	ATTGGTCACC	GGAGAAAGGC	CACACCTGTT	GTGGGAAAAA		1250
GCAAAAGAAC	AACGAAACTT	GTGGTGAAGC	AGGTGAAAAC	GGAGTTTCCA		1300
AGAAAAGTCC	TGTTAACTAT	GGAAGGATGA	TATCTTTTGG	GAAAGAAGTA		1350
GCAGAGGCTG	CGCCATCTGA	GATTAATAAT	ATTGATTTTC	GAGGTGCTGT		1400
CAAAGGTCAG	AGTATCCCAA	ATCACACCTG	TCGTGACGTG	TGGACAGAGT		1450
ACCATGACAT	GGGAATTGCT		_	GTATAAGGTC		1500
TACACTGCTG	GTGAAGCTAT			CTCCTAAGAT		1550
GATGGCGCGT	GGTGCCGCTC	ATTTCTCTTA	TGGAATTGCT	GATGATTTGG		1600
ATGACACCAA	. GTATCAAGAT	CCCAAATACT	GGTCAAATCC	GTTAGAGACA		1650
AAATTACCGA	ATGCTCCTGA	. GATGGAAATC	TACTCATTAT	ACGGAGTGGG		1700
	GAACGAGCAT			TCTCCCGACA		1750
GTTGCATCCC	CTTTCAGATA			GGACGAAGAT		1800
AGCTGTCTGA	AAGCAGGAGT			AAACAGTACC		1850
CGTCCTAAGI				GGCAAGACAA		1900
GATTCAACCC		: AAGACTTATA		CAATCACTCT		1950
			ACGCAGAGTG			2000
			AGATATCATG			2050
				CTCTGGCATA		2100
				CATGATCTCT		2150
				AAGAGAGATC		2200
				CTCACAAAGA		2250
				CTTTAATATT		2300
				ACGATAACAC		2350
			TGTCAAAAGC	ATCAATTTGT		2400
GGGTTAAAA	AAAAAAAA	AAAAAA		2427		







<210> 6				<del></del>	,
<211> 671					
<212> PRT					
<213> Arabidopsis th	naliana				
<400> 6					
MPLIHRKKPT EKPSTPPS	EE VVHDEDSQKK			50	
IDSCCWFIGC VCVTWWFL	LF LYNAMPASFP	QYVTERITGP	LPDPPGVKLK	100	
KEGLKAKHPV VFIPGIVT	GG LELWEGKQCA	DGLFRKRLWG	GTFGEVYKRP	150	
LCWVEHMSLD NETGLDPA	GI RVRAVSGLVA	ADYFAPGYFV	WAVLIANLAH	200	
IGYEEKNMYM AAYDWRLS	FQ NTEVRDQTLS	RMKSNIELMV	STNGGKKAVI	250	
VPHSMGVLYF LHFMKWVE	AP APLGGGGGPD	WCAKYIKAVM	NIGGPFLGVP	300	
KAVAGLFSAE AKDVAVAR	AI APGFLDTDIF	RLQTLQHVMR	MTRTWDSTMS	350	
MLPKGGDTIW GGLDWSPE	KG HTCCGKKQKN	NETCGEAGEN	GVSKKSPVNY	400	
GRMISFGKEV AEAAPSEI	NN IDFRGAVKGQ	SIPNHTCRDV	WTEYHDMGIA	450	
GIKAIAEYKV YTAGEAID	LL HYVAPKMMAR	GAAHFSYGIA	DDLDDTKYQD	500	
PKYWSNPLET KLPNAPEM	EI YSLYGVGIPT	ERAYVYKLNQ	SPDSCIPFQI	550	
FTSAHEEDED SCLKAGVY	NV DGDETVPVLS	AGYMCAKAWR	GKTRFNPSGI	600	
KTYIREYNHS PPANLLEG	RG TOSGAHVDIM	GNFALIEDIM	RVAAGGNGSD	650	
IGHDOVHSGI FEWSERIL	<del></del>	671			



<210> 7 <211> 643 <212> cDNA <213> Zea mays

<221> CDS <222> (1)..(402)

<400> 6
CGG GAG AAA ATA GCT GCT TTG AAG GGG GGT GTT TAC TTA GCC GAT GGT 48
Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly
1 5 10 15

GAT GAA ACT GTT CCA GTT CTT AGT GCG GGC TAC ATG TGT GCG AAA GGA 96
Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly
20 25 30

TGG CGT GGC AAA ACT CGT TTC AGC CCT GCC GGC AGC AAG ACT TAC GTG 144
Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val
35 40 45

AGA GAA TAC AGC CAT TCG CCA CCC TCT ACT CTC CTG GAA GGC AGG GGC 192
Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly
50 55 60

ACC CAG AGC GGT GCA CAT GTT GAT ATA ATG GGG AAC TTT GCT CTA ATT 240 Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile 65 70 75 80

GAG GAC GTC ATC AGA ATA GCT GCT GGG GCA ACC GGT GAG GAA ATT GGT 288
Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly
85 90 95

GGC GAT CAG GTT TAT TCA GAT ATA TTC AAG TGG TCA GAG AAA ATC AAA 336 Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys 100 105 110

TTG AAA TTG TAA CCTATGGGAA GTTAAAGAAG TGCCGACCCG TTTATTGCGTTCC 391 Leu Lys Leu

AAAGTGTCCT GCCTGAGTGC AACTCTGGAT TTTGCTTAAA TATTGTAATT TTTCACGC 449
TTCATTCGTC CCTTTGTCAA ATTTACATTT GACAGGACGC CAATGCGATA CGATGTTG 507
TACCGCTATT TTCAGCATTG TATATTAAAC TGTACAGGTG TAAGTTGCAT TTGCCAGC 565
TGAAATTGTG TAGTCGTTTT CTTTACGATT TAATANCAAG TGGCGGAGCA GTGCCCCA 623
AGCNAAAAAA AAAAAAAAA



<210>8

<211> 115

<212> PRT

<213> Zea mays

<400> 8

Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly 1 5 10 15

Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly 20 25 30

Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val 35 40 45

Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly 50 60

Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile 65 70 75 80

Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly 85 90 95

Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys 100 105 110

Leu Lys Leu 115



<400> 9 ggtggcgaag acganggcgg aagttggagg ctaacgagaa tgacnctcgg 50 agatggatct accetctaga gacacgacta centtgcace cagectcaag 100 gtntacngtt tntatgggta ggaagcegac ggagcgagce tacatctate 150 tggcgcccga tcccgggacg acaacgcate tttagatgac gatcgatacg 200 actttgactn aggggcacat tgaccacggt gtgattttgg gcgaaggcga 250 tggcacagtg aacettatga gtttggggta cetgtgcaat aaggggtgga 300 aaatgaagag atacaatcct gcgggetcaa aaataaccgt ggtcgagatg 350 ccgcatgaac cagaacggtt caatccgaga ggagggccga atacggcgga 400 tcacgtggat attetaggaa ggcagaatct aaacgagtac attettaaag 450 tggcggcagg tcgaggcgat acaattgagg attttattac tagtaatatt 500 cttaaatatg tagaaaaggt tgaaatttat gaagagtaat taaatacggc 550 acataggtta ctcaatagta tgactaatta aaaaaaaatt ttttttctaa, 600 aaaaaaaaaaa aaaaa 616	<210> 9 <211> 616 <212> cDNA <213> Neuro	espora crass			,	
ggtggcgaag acganggcgg aagttggagg ctaacgagaa tgacnctcgg 50 agatggatct accetctaga gacacgacta centtgcace cagectcaag 100 gtntacngtt tntatgggta ggaagccgac ggagcgagce tacatctate 150 tggcgccga tecegggacg acaacgcate tttagatgac gategatacg 200 actttgactn aggggcacat tgaccacggt gtgattttgg gegaaggega 250 tggcacagtg aacettatga gtttggggta cetgtgcaat aaggggtgga 300 aaatgaagag atacaatect gegggetcaa aaataacegt ggtegagatg 350 cegcatgaac cagaacggtt caateegaga ggagggeega atacggegga 400 tcacgtggat attetaggaa ggcagaatet aaacgagtac attettaaag 450 tggeggeagg tegaggegat acaattgagg attttattac tagtaatatt 500 cttaaatatg tagaaaaggt tgaaatttat gaagagtaat taaatacgge 550 acataggtta etcaatagta tgactaatta aaaaaaaaatt tttttecaa. 600	12257 110410	Spora crass	, ω			
agatggatct accetctaga gacacgacta centtgcace cagectcaag gtntacngtt tntatgggta ggaagecgae ggagegagee tacatetate 150 tggegeega teeegggaeg acaacgeate tttagatgae gategataeg 200 actttgactn aggggeacat tgaceaeggt gtgattttgg gegaaggega 250 tggeaeagtg aacettatga gtttggggta cetgtgeaat aaggggtgga 300 aaatgaagag atacaateet gegggetcaa aaataaeegt ggtegagatg 350 cegeatgaae cagaaeggtt caateegaga ggagggeega atacggegga 400 teaegtggat attetaggaa ggeagaatet aaaegagtae attettaaag 450 tggeggeagg tegaggegat acaattgagg attttattae tagtaatatt 500 cetaaatatg tagaaaaggt tgaaatttat gaagagtaat taaataegge 550 acataggtta etcaatagta tgaetaatta aaaaaaaaatt tttttetaa.	<400> 9					
gtntacngtt tntatgggta ggaagccgac ggagcgagcc tacatctatc 150 tggcgccga tcccgggacg acaacgcatc tttagatgac gatcgatacg 200 actttgactn aggggcacat tgaccacggt gtgattttgg gcgaaggcga 250 tggcacagtg acccttatga gtttggggta cctgtgcaat aaggggtgga 300 aaatgaagag atacaatcct gcgggctcaa aaataaccgt ggtcgagatg 350 ccgcatgaac cagaacggtt caatccgaga ggagggccga atacggcgga 400 tcacgtggat attctaggaa ggcagaatct aaacgagtac attcttaaag 450 tggcggcagg tcgagggct acaattgagg attttattac tagtaatatt 500 cttaaatatg tagaaaaggt tgaaatttat gaagagtaat taaatacggc 550 acataggtta ctcaatagta tgactaatta aaaaaaaaatt ttttttctaa. 600	ggtggcgaag	acganggcgg	aagttggagg	ctaacgagaa	tgacnctcgg	50
tggcgccga tcccgggacg acaacgcatc tttagatgac gatcgatacg 200 actttgactn aggggcacat tgaccacggt gtgattttgg gcgaaggcga 250 tggcacagtg aaccttatga gtttggggta cctgtgcaat aaggggtgga 300 aaatgaagag atacaatcct gcgggctcaa aaataaccgt ggtcgagatg 350 ccgcatgaac cagaacggtt caatccgaga ggagggccga atacggcgga 400 tcacgtggat attctaggaa ggcagaatct aaacgagtac attcttaaag 450 tggcggcagg tcgaggcgat acaattgagg attttattac tagtaatatt 500 cttaaatatg tagaaaaggt tgaaatttat gaagagtaat taaatacggc 550 acataggtta ctcaatagta tgactaatta aaaaaaaaatt ttttttctaa. 600	agatggatct	accctctaga	gacacgacta	ccnttgcacc	cagcctcaag	100
actttgactn aggggcacat tgaccacggt gtgattttgg gcgaaggcga 250 tggcacagtg aaccttatga gtttggggta cctgtgcaat aaggggtgga 300 aaatgaagag atacaatcct gcgggctcaa aaataaccgt ggtcgagatg 350 ccgcatgaac cagaacggtt caatccgaga ggagggccga atacggcgga 400 tcacgtggat attctaggaa ggcagaatct aaacgagtac attcttaaag 450 tggcggcagg tcgaggcgat acaattgagg attttattac tagtaatatt 500 cttaaatatg tagaaaaggt tgaaatttat gaagagtaat taaatacggc 550 acataggtta ctcaatagta tgactaatta aaaaaaaaatt ttttttctaa. 600	gtntacngtt	tntatgggta	ggaagccgac	ggagcgagcc	tacatctatc	150
tggcacagtg aaccttatga gtttggggta cctgtgcaat aaggggtgga 300 aaatgaagag atacaatcct gcgggctcaa aaataaccgt ggtcgagatg 350 ccgcatgaac cagaacggtt caatccgaga ggagggccga atacggcgga 400 tcacgtggat attctaggaa ggcagaatct aaacgagtac attcttaaag 450 tggcggcagg tcgaggcgat acaattgagg attttattac tagtaatatt 500 cttaaatatg tagaaaaggt tgaaatttat gaagagtaat taaatacggc 550 acataggtta ctcaatagta tgactaatta aaaaaaaaatt ttttttctaa. 600	tggcgcccga	tcccgggacg	acaacgcatc	tttagatgac	gatcgatacg	200
aaatgaagag atacaatcct gcgggctcaa aaataaccgt ggtcgagatg 350 ccgcatgaac cagaacggtt caatccgaga ggagggccga atacggcgga 400 tcacgtggat attctaggaa ggcagaatct aaacgagtac attcttaaag 450 tggcggcagg tcgaggcgat acaattgagg attttattac tagtaatatt 500 cttaaatatg tagaaaaggt tgaaatttat gaagagtaat taaatacggc 550 acataggtta ctcaatagta tgactaatta aaaaaaaaatt ttttttctaa, 600	actttgactn	aggggcacat	tgaccacggt	gtgattttgg	gcgaaggcga	250
ccgcatgaac cagaacggtt caatccgaga ggagggccga atacggcgga 400 tcacgtggat attctaggaa ggcagaatct aaacgagtac attcttaaag 450 tggcggcagg tcgaggcgat acaattgagg attttattac tagtaatatt 500 cttaaatatg tagaaaaggt tgaaatttat gaagagtaat taaatacggc 550 acataggtta ctcaatagta tgactaatta aaaaaaaaatt ttttttctaa, 600	tggcacagtg	aaccttatga	gtttggggta	cctgtgcaat	aaggggtgga	300
tcacgtggat attctaggaa ggcagaatct aaacgagtac attcttaaag 450 tggcggcagg tcgaggcgat acaattgagg attttattac tagtaatatt 500 cttaaatatg tagaaaaggt tgaaatttat gaagagtaat taaatacggc 550 acataggtta ctcaatagta tgactaatta aaaaaaaaatt ttttttctaa. 600	aaatgaagag	atacaatcct	gcgggctcaa	aaataaccgt	ggtcgagatg	350
tggcggcagg tcgaggcgat acaattgagg attttattac tagtaatatt 500 cttaaatatg tagaaaaggt tgaaatttat gaagagtaat taaatacggc 550 acataggtta ctcaatagta tgactaatta aaaaaaaaatt ttttttctaa, 600	ccgcatgaac	cagaacggtt	caatccgaga	ggagggccga	atacggcgga	400
cttaaatatg tagaaaaggt tgaaatttat gaagagtaat taaatacggc 550 acataggtta ctcaatagta tgactaatta aaaaaaaaatt ttttttctaa, 600	tcacgtggat	attctaggaa	ggcagaatct	aaacgagtac	attcttaaag	450
acataggtta ctcaatagta tgactaatta aaaaaaaatt ttttttctaa, 600	tggcggcagg	tcgaggcgat	acaattgagg	attttattac	tagtaatatt	500
	cttaaatatg	tagaaaaggt	tgaaatttat	gaagagtaat	taaatacggc	550
aaaaaaaaa aaaaaa 616	acataggtta	ctcaatagta	tgactaatta	aaaaaaaatt	ttttttctaa,	600
	aaaaaaaaa	aaaaaa		616	5	



GTCAATGAAT GA





.01/0	
<210> 10	
<211> 1562	
<212> genomic DNA	
<213> Arabidopsis thaliana	
•	
<400> 10	
ATGAAAAAA TATCTTCACA TTATTCGGTA GTCATAGCGA TACTCGTTGT	50
GGTGACGATG ACCTCGATGT GTCAAGCTGT GGGTAGCAAC GTGTACCCTT	100
TGATTCTGGT TCCAGGAAAC GGAGGTAACC AGCTAGAGGT ACGGCTGGAC	150
AGAGAATACA AGCCAAGTAG TGTCTGGTGT AGCAGCTGGT TATATCCGAT	200
TCATAAGAAG AGTGGTGGAT GGTTTAGGCT ATGGTTCGAT GCAGCAGTGT	250
TATTGTCTCC CTTCACCAGG TGCTTCAGCG ATCGAATGAT GTTGTACTAT	300
GACCCTGATT TGGATGATTA CCAAAATGCT CCTGGTGTCC AAACCCGGGT	350
TCCTCATTTC GGTTCGACCA AATCACTTCT ATACCTCGAC CCTCGTCTCC	400
GGTTAGTACT TTCCAAGATA TATCATTTTG GGACATTTGC ATAATGAACA	450
AAATAGACAT AAATTTGGGG GATTATTGTT ATATCAATAT CCATTTATAT	500
GCTAGTCGGT AATGTGAGTG TTATGTTAGT ATAGTTAATG TGAGTGTTAT	550
GTGATTTTCC ATTTTAAATG AAGCTAGAAA GTTGTCGTTT AATAATGTTG	600
CTATGTCATG AGAATTATAA GGACACTATG TAAATGTAGC TTAATAATAA	650
GGTTTGATTT GCAGAGATGC CACATCTTAC ATGGAACATT TGGTGAAAGC	700
TCTAGAGAAA AAATGCGGGT ATGTTAACGA CCAAACCATC CTAGGAGCTC	750
CATATGATTT CAGGTACGGC CTGGCTGCTT CGGGCCACCC GTCCCGTGTA	800
GCCTCACAGT TCCTACAAGA CCTCAAACAA TTGGTGGAAA AAACTAGCAG	850
CGAGAACGAA GGAAAGCCAG TGATACTCCT CTCCCATAGC CTAGGAGGAC	900
TTTTCGTCCT CCATTTCCTC AACCGTACCA CCCCTTCATG GCGCCGCAAG	950
TACATCAAAC ACTTTGTTGC ACTCGCTGCG CCATGGGGTG GGACGATCTC	1000
TCAGATGAAG ACATTTGCTT CTGGCAACAC ACTCGGTGTC CCTTTAGTTA	1050
ACCCTTTGCT GGTCAGACGG CATCAGAGGA CCTCCGAGAG TAACCAATGG	1100
CTACTTCCAT CTACCAAAGT GTTTCACGAC AGAACTAAAC CGCTTGTCGT	1150
AACTCCCCAG GTTAACTACA CAGCTTACGA GATGGATCGG TTTTTTGCAG	1200
ACATTGGATT CTCACAAGGA GTTGTGCCTT ACAAGACAAG	1250
TTAACAGAGG AGCTGATGAC TCCGGGAGTG CCAGTCACTT GCATATATGG	1300
GAGAGGAGTT GATACACCGG AGGTTTTGAT GTATGGAAAA GGAGGATTCG	1350
ATAAGCAACC AGAGATTAAG TATGGAGATG GAGATGGGAC GGTTAATTTG	1400
GCGAGCTTAG CAGCTTTGAA AGTCGATAGC TTGAACACCG TAGAGATTGA	1450
TGGAGTTTCG CATACATCTA TACTTAAAGA CGAGATCGCA CTTAAAGAGA	1500
TTATGAAGCA GATTTCAATT ATTAATTATG AATTAGCCAA TGTTAATGCC	1550

1562



<210> 11 <211> 3896 <212> genomic DNA <213> Arabidopsis thaliana <400> 11 ATGGGAGCGA ATTCGAAATC AGTAACGGCT TCCTTCACCG TCATCGCCGT 50 TTTTTTTTG ATTTGCGGTG GCCGAACTGC GGTGGAGGAT GAGACCGAGT 100 TTCACGGCGA CTACTCGAAG CTATCGGGTA TAATCATTCC GGGATTTGCG 150 TCGACGCAGC TACGAGCGTG GTCGATCCTT GACTGTCCAT ACACTCCGTT 200 GGACTTCAAT CCGCTCGACC TCGTATGGCT AGACACCACT AAGGTCCGTG 250 300 350 400 450 TAGCGCAATG AATGTGTAAT TAGTCTGCGC TTTTTATTCA ACTAGATCTG 500 CAAGTTTTTC AGAGTGCTCA ATAGTAGTTA GAAAATGTTA GGTCATTTTA 550 CTTGTGCATT GTGATTCTTT TGGTTGTTGC TTACTGATCG ACGTGATGGA 600 TGGTTTACAG CTTCTTTCTG CTGTCAACTG CTGGTTTAAG TGTATGGTGC 650 TAGATCCTTA TAATCAAACA GACCATCCCG AGTGTAAGTC ACGGCCTGAC 700 AGTGGTCTTT CAGCCATCAC AGAATTGGAT CCAGGTTACA TAACAGGTAG 750 TTTCGGATTT TTCTTTCTTT TGAGTTTTCT TCAATTTGAT ATCATCTTGT 800 TGTGATATAA TATGGCTAAG TTCATTAATT TGGTCAATTT TCAGGTCCTC 850 TTTCTACTGT CTGGAAAGAG TGGCTTAAGT GGTGTGTTGA GTTTGGTATA 900 GAAGCAAATG CAATTGTCGC TGTTCCATAC GATTGGAGAT TGTCACCAAC 950 CAAATTGGAA GAGCGTGACC TTTACTTTCA CAAGCTCAAG TTAGTCCTTA 1000 TCAGGCTAAT GTCTTTTATC TTCTCTTTTT ATGTAAGATA AGCTAAGAGC 1050 TCTGGTCGTC TTCCTTTTTG CAGGTTGACC TTTGAAACTG CTTTAAAACT CCGTGGCGGC CCTTCTATAG TATTTGCCCA TTCAATGGGT AATAATGTCT TCAGATACTT TCTGGAATGG CTGAGGCTAG AAATTGCACC AAAACATTAT TTGAAGTGGC TTGATCAGCA TATCCATGCT TATTTCGCTG TTGGTACCGG CCTACTATCC TTAAGTTACC ATTTTATTTT TTCTCTAATT GGGGGAGTTA TGTTGTGACT TACTGGATTG AGCTCGATAC CTGATTTGTT GTTGATTTAG GAGCTCCTCT TCTTGGTTCT GTTGAGGCAA TCAAATCTAC TCTCTCTGGT GTAACGTTTG GCCTTCCTGT TTCTGAGGTG ACCTCTGACT TCTCTTTAGT 1450 TTTAAGTAGT TGATATCAAC CAGGTCTTAT AACTCACTGG ATTTTCCTTT 1500 TGAAAGTATT ACTTTTGTTA ATTGAACTGC TGTACGCGAT ATGGTATCTG 1550 TAGATCTTGA AGTGCTAGTT ATCAAAGAAC ATATTGTGGG TAGTATACCT 1600 GTCAGCGGCC TTAGCTAATA CAACCAAACC ACATGTACAC TGATTTAGTT 1650 TTCAGATTAT TATGGTAGAC TTTAAGTTGA GAAGAACTT TGACTGAAAT 1700 CTTTTTATTT TAATAGGCTA TGATTTGTTT ATTGAAATCA TGTGACATAT 1750 TGACATGCGC TTCTCATGTT TTTTGTTGGC AAGGCTTCAG GGAACTGCTC 1800 GGTTGTTGTC CAATTCTTTT GCGTCGTCAT TGTGGCTTAT GCCATTTTCA 1850 AAGAATTGCA AGGGTGATAA CACATTCTGG ACGCATTTTT CTGGGGGTGC 1900 TGCAAAGAAA GATAAGCGCG TATACCACTG TGATGAAGAG GAATATCAAT 1950 CAAAATATTC TGGCTGGCCG ACAAATATTA TTAACATTGA AATTCCTTCC 2000 ACTAGCGGTT AGACTCTGTA TATGCAACTG TAACACTAAC AAAAGTTTCA 2050 CCAAGAATGT TCACTCTCAT ATTTCGTTCC TTTGATGTGT ATCCATCAGT 2100 TACAGAAACA GCTCTAGTCA ACATGACCAG CATGGAATGT GGCCTTCCCA 2150 CCCTTTTGTC TTTCACAGCC CGTGAACTAG CAGATGGGAC TCTTTTCAAA 2200 GCAATAGAAG ACTATGACCC AGATAGCAAG AGGATGTTAC ACCAGTTAAA 2250 GAAGTACGTA CCTTTCTTTG TGATAAGAAA TATTGCTCAT CGATCATCAC 2300 TTGCTGGCTT CTTGTACGTC AAATTGTTTT GTTTAAATCT CTATATCAAT 2350 TGTTCATATG CTTTGTCTTT CTTACTATAA GAAACAAGTA TAATCAGAAA 2400 CCTTATTATT GATTATCAGT TCTCTCCTTA TATTATGGAA TGTCTTTTTC 2450 GTTTACAGTT ATGAATGCAA AAGGGGGTAT TTTAGTTGAT TGATTCTCTC 2500 ATTCTCTAGT TTGTTTTGAC TAATAGCGTC AATTTTGTTT TTCTAGCAAA 2550 TCTTTGTGAA TTATATATAA CATGCTAACT ATACTTTTCA GGTTGTATCA 2600 TGATGACCCT GTTTTTAATC CTCTGACTCC TTGGGAGAGA CCACCTATAA 2650 AAAATGTATT TTGCATATAT GGTGCTCATC TAAAGACAGA GGTATGATGC 2700 ATTCTCAATA TCACATTATG CGTTGACTTT GTTATTATAT TCCCCATTTG 2750 GTTTGCAATA TCTTTTTGAA TTATGATTTA TCTTCTCCCT TGCATCTTAT 2800 GCTATTAAGC GTTAAAGGTA CTAAATGTAT GAAGCTGTCT GTCATAGGTT





3896

WO 00/60095	5		17 / 53	TC1/I	2P00/02701	
GGTTATTACT	TTGCCCCAAG	TGGCAAACCT	TATCCTGATA	ATTGGATCAT		<del>2900</del>
CACGGATATC	ATTTATGAAA	CTGAAGGTTC	CCTCGTGTCA	AGGTAATTTT		2950
CCGCAATGGC	AGAAGTAAAA	CAGGAAGGCA	AAGTCTTCTG	TATCAGTCTA		3000
GTGGCATGTT	ATCTCAGTTG	CATAAGCAAA	TTATTAAACA	ACTAAAATTT		3050
AAGTACTTTT	TTATCATTCC	TTTTGAGCTT	AGTGGATGAT	CAGTGGCTTA		3100
AAGTGGGAAG	AGGTGTTGCA	TGAAACATGA	CACTTGTATC	AAAGATAACT		3150
AGCAAAACAA	AACTAACCCA	TTTCTGAATT	TCATATTATT	AGGAGTAGTC		3200
GTGCTTTTAA	AAAATTTGTT	TTAAGAAACC	GAAAAACTAG	TTCATATCTT		3250
GATTGTGCAA	TATCTGCAGG	TCTGGAACTG	TGGTTGATGG	GAACGCTGGA	3300	
CCTATAACTG	GGGATGAGAC	GGTAAGCTCA	GAAGTTGGTT	TTGAAATTAT		3350
CTTCTTGCAA	ACTACTGAAG	ACTAAGATAA	TACTTGCTTC	TGGAACACTG		3400
CTTGCTATGT	TCTCTAGTAC	ACTGCAATAT	TGACTCTCCG	CTACTTTTAT		3450
TGATTATGAA	ATTGATCTCT	TATAGGTACC	CTATCATTCA	CTCTCTTGGT		3500
GCAAGAATTG	GCTCGGACCT	AAAGTTAACA	TAACAATGGC	TCCCCAGGTA		3550
CTCTTTTTTA	GTTCCTCACC	TTATATAGAT	CAAACTTTAA	GTGTACTTTT		3600
CTGGTTATGT	GTTGATTTAC	CTCCAATTTG	TTCTTTCTAA	AAATCATATA		3650
TCTCTGTACT	CCTCAAGAAC	TTGTATTAAT	CTAAACGAGA	TTCTCATTGG		3700
GAAAATAAAA	CAACAGCCAG	AACACGATGG	AAGCGACGTA	CATGTGGAAC		3750
TAAATGTTGA	TCATGAGCAT	GGGTCAGACA	TCATAGCTAA	CATGACAAAA		3800
GCACCAAGGG	TTAAGTACAT	AACCTTTTAT	GAAGACTCTG	AGAGCATTCC		3850

GGGGAAGAGA ACCGCAGTCT GGGAGCTTGA TAAAAGTGGG TATTAA



<210> 12 <211> 709 <212> cDNA <213> Lycopersicon escul	entum			
<400> 12				
CTGGGGCCAA AAGTGAACAT	AACAAGGACA	CCACAGTCAG	AGCATGATGT	50
TCAGATGTAC AAGTGCATCT	AAATATAGAG	CATCAACATG	GTGAAGATAT	100
CATTCCCAAT ATGACAAAGT	TACCTACAAT	GAAGTACATA	ACCTATTATG	150
AGGATTCTGA AAGTTTTCCA	GGGACAAGAA	CAGCAGTTTG	GGAGCTTGAT	200
AAAGCAAATC ACAGGAACAT	TGTCAGATCT	CCAGCTTTGA	TGCGGGAGCT	250
GTGGCTTGAG ATGTGGCATG	ATATTCATCC	TGATAAAAAG	TCCAAGTTTG	300
TTACAAAAGG TGGTGTCTGA	TCCTCACTAT	TTTCTTCTAT	AAATGTTTGA	350
GTTTGTATTG ACATTGTAAG	TATTGCAACA	AAAAGCAAAG	CGTGGGCCTC	400
TGAGGGATGA GGACTGCTAT	TGGGATTACG	GGAAAGCTCG	ATGTGCATGG	450
GCTGAACATT GTGAATACAG	GTTAGAATAT	TCAAATTATA	TTTTGCAAAA	500
TATTCTCTTT TTGTGTATTT	AGGCCACCTT	TCCCCGGTCA	CAACGATGCA	550
GATATGTATT CGGGGATGTT	CACCTGGGAC	AGAGTTGCAG	ATTGAAGAGT	600
TCTACATCTC ACATCCTGTC	ACACTATGTG	TGATATTTAA	GAAACTTTGT	650
TTGGCGGAAC AACAAGTTTG	CACAAACATT	TGAAGAAGAA	AGCGAAATGA	700
TTCAGAGAG		709	)	

#### WO 00/60095

19 / 53



<210> 13

<211> 623

<212> PRT

<213> Schizosaccharomyces pombe

<400> 13

MASSKKSKTHKKKKEVKSPIDLPNSKKPTRALSEQPSASETQSVSNKSRKSKFGKRLNFILGAILGICGA70
FFFAVGDDNAVFDPATLDKFGNMLGSSDLFDDIKGYLSYNVFKDAPFTTDKPSQSPSGNEVQVGLDMYNE140
GYRSDHPVIMVPGVISSGLESWSFNNCSIPYFRKRLWGSWSMLKAMFLDKQCWLEHLMLDKKTGLDPKGI210
KLRAAQGFEAADFFITGYWIWSKVIENLAAIGYEPNNMLSASYDWRLSYANLEERDKYFSKLKMFIEYSN280
IVHKKKVVLISHSMGSQVTYYFFKWVEAEGYGNGGPTWVNDHIEAFINISGSLIGAPKTVAALLSGEMKD350
TGIVITLNILEKFFSRSERAMMVRTMGGVSSMLPKGGDVAPDDLNQTNFSNGAIIRYREDIDKDHDEFDI420
DDALQFLKNVTDDDFKVMLAKNYSHGLAWTEKEVLKNNEMPSKWINPLETSLPYAPDMKIYCVHGVGKPT490
ERGYYYTNNPEGQPVIDSSVNDGTKVENGIVMDDGDGTLPILALGLVCNKVWQTKRFNPANTSITNYEIK560
HEPAAFDLRGGPRSAEHVDILGHSELNEIILKVSSGHGDSVPNRYISDIQEIINEINLDKPRN 623

#### WO 00/60095

20 / 53



<210> 14

<211> 432

<212> PRT

<213> Arabidopsis thaliana

<400> 14
MKKISSHYSVVIAILVVVTMTSMCQAVGSNVYPLILVPGNGGNQLEVRLDREYKPSSVWCSSWLYPIHKK70
SGGWFRLWFDAAVLLSPFTRCFSDRMMLYYDPDLDDYQNAPGVQTRVPHFGSTKSLLYLDPRLRDATSYM140
EHLVKALEKKCGYVNDQTILGAPYDFRYGLAASGHPSRVASQFLQDLKQLVEKTSSENEGKPVILLSHSL210
GGLFVLHFLNRTTPSWRRKYIKHFVALAAPWGGTISQMKTFASGNTLGVPLVNPLLVRRHQRTSESNQWL280
LPSTKVFHDRTKPLVVTPQVNYTAYEMDRFFADIGFSQGVVPYKTRVLPLTEELMTPGVPVTCIYGRGVD350
TPEVLMYGKGGFDKQPEIKYGDGDGTVNLASLAALKVDSLNTVEIDGVSHTSILKDEIALKEIMKQISII420
NYELANVNAVNE



<210> 15

<211> 552

<212> PRT

<213> Arabidopsis thaliana

<400> 15

MGANSKSVTASFTVIAVFFLICGGRTAVEDETEFHGDYSKLSGIIIPGFASTQLRAWSILDCPYTPLDFN70
PLDLVWLDTTKLLSAVNCWFKCMVLDPYNQTDHPECKSRPDSGLSAITELDPGYITGPLSTVWKEWLKWC140
VEFGIEANAIVAVPYDWRLSPTKLEERDLYFHKLKLTFETALKLRGGPSIVFAHSMGNNVFRYFLEWLRL210
EIAPFHYLKWLDQHIHAYFAVGAPLLGSVEAIKSTLSGVTFGLPVSEGTARLLSNSFASSLWLMPFSKNC280
KGCNTFWTHFSGGAAKKDKRVYHCDEEEYQSKYSGWPTNIINIEIPSTSARELADGTLFKAIEDYDPDSK350
RMLHQLKKYVPFFVIRNIAHRSSLAGFLLYHDDPVFNPLTPWERPPIKNVFCIYGAHLKTEVGYYFAPSG420
KPYPDNWIITDIIYETEGSLVSRSGTVVDGNAGPITGDETVPYHSLSWCKNWLGPKVNITMAPQILIGKI490
KQQPEHDGSDVHVELNVDHEHGSDIIANMTKAPRVKYITFYEDSESIPGKRTAVWELDKSGY
552

5

10

15 <170> PatentIn Ver. 2.0 <210> 10 <211> 661 20 <212> PRT <213> Saccharomyces cerevisiae <400> 1 Mer Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser 25 Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg 20 25 30 Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Gly 35 40 30 Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg 50 55 60 35 Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu 65 70 75 80 Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe 85 90 95 40 Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe 100 105 110 Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val 115 120 125 45 Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn 130 135. 50 Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly 145 150 150 Lys Glm Leu Leu Arg Asp Tyr Asm Ile Glu Ala Lys His Pro Val Val 165 170 175 55 Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile 180 185 190 Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp 195 200 205 60

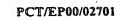
Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp
5 Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn 240 225
Phe Thr Leu Arg Ala Ala Glr Gly Phe Glu Ser Thr Asp Tyr Phe Ile 255
O Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile 265
Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu 285 275
Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys 290 290
20 Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu 320 315 305
Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp 335 325
Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val 340 340 340 340 340 340 340 340 340
Asn Glu Kis Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly 365 30 355
Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr 370 380
35 Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser 400 395 385 390 390 397 400 398
Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser 410 405 40 Gly Asp Met Lys Ser Ser
Met Leu Pro Lys Gly Glu Glu Val 115 119 430
Ser Glu Asp Ala Leu Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile 445 45 45 48 48 48 49 49 40 40 40 40 40 40 40 40 40
Arg Phe Glu Arg Ash Thr Ser Asp Ala Phe Ash Lys Ash Leu Thr Met 450 450 450 450 450 450 450 460 450 450 460
50 Lys Asp Ala Ile Asm Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln 480 475 465
Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu Glu 490 495
Glu Leu Arg Lys Ash Glu Leu His his 250 510
Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr 60 515 520

	Gly	Val	. As	n i	Asn	Pro	Thr	Glu 535	Arg	Ala	r TY:	. Val	Ty::	Lys	Glu	: Glu	ı Ası	Þ
5	545						220					555 555						
						56:	•				٠.							
10					580	)					_	r Pro						
			5	95					00	U		s Gl						
15		63	10					0.1	. >			Ls Va						
20	52	5					ده	U				le Al 63						
	Ľe	in A	ai (	Glu	: Pr	o A1	rg Gl 15	n Le	eu Se	er A	sn L 6	eu Se 50	er G	in T	۸ مت	al 5 6	er G 55	±n
25	Me	et P	TO	Phe	6 P =	C M	et											
30	< <	210> 211> 212: 213:	3 E •	37 2T	ido	psis	tha	liar	ıa									
35	V	400 al 1	> 2 Gly	Se	er A	sn V	/al T 5	yr !	Pro I	Leu	Ile :	Leu V 10	al :	?zo (	ely A	Asn (	Gly 15	GļĀ
						20					2.2	Glu <sup>1</sup>						
40	,	طت	Cys	s Sé	er 9	er	Trp !	Leu	ŢYI	Pro 40	Ile	His :	Lys '	Lys	Ser 45	Gly	GīĀ	طتت
45	5	Phe	A≃g 50	; L	eu :	LTD	Phe	ÇzA	Ala 55	Ala	Val	Leu	Leu	S <b>e</b> ≍ 50	PTO	Phe	Thr	ATG
		Cys 55	274	e S	er .	Asp	Αrç	Met 70	Met	Leu	፲አ <del>፲</del>	īāī	A.S.D 75	Pro	Asp	ren	Asp	QzA 08
5	0	Tyr	Gl:	n A	ısn	Ala	P=0 85	GŗĀ	Val	Gln	تتت	Arg 90	Val	P≖o	Kis	Phe	Gly 95	Ser
		Tin	īy	s S	er	Leu 100	Leu	سَنَّت	Leu	Asp	PT0 105	Arg	Leu	Arg	Asp	Ala 110	The	Ser
5	5	Ty=	Me	:	31u 115	His	Leu	Val	Γλε	Al= 120	Leu	Glu	Lys	lys	Cys 125	Gly	ፊላቷ	val
,	50	AS:	: As	; c:	Gln	The	Ile	Гел	Gly 135	Ala	. Pro	Tyr	Asp	Phe 140	A≝g	. Lâr	Gly	/ Leu

	Ala 1	Ala	Ser	Gly	His	Pro 150	Ser	Arg	Val	Ala	Ser 155	Gln	Phe	Leu	Gln	Asp 160
5	Leu !	Lys	Gln	Leu	Val 165	Glu	ŗās	Thr	Ser	Ser 170	Glu	Asn	Glu	Gly	Lys 175	Pro
	Val	Ile	Leu	Leu 180	Ser	His	Ser	Leu	Gly 185	Gly	Leu	Phe	Val	Leu 190	His	Phe
10	Leu	Asn	Arg 195	Thr	Thr	Pro	Ser	Trp 200	Arg	Arg	Lys	Tyr	11e 205	Lys	His	Phe
, -	Val	Ala 210	Leu	Ala	Ala	Pro	T:p 215	Gly	Gly	Thr	Ile	Ser 220	Gln	Met	Lys	Thr
15	Phe 225	Ala	Ser	Gly	Asn	Thr 230	Leu	Gly	Val	Pro	Leu 235	Val	Asn	Pro	Leu	Leu 240
20	Val	Arg	Arg	His	Gln 245	Arg	Thr	Ser	Glu	Se:	Asn )	Gln	Trp	Leu	Leu 255	Pro
	Ser	Thr	. Tās	Val 260	Phe	His	Asp	ATS	7hr 265	Lys	s Pro	ren	Val	Val 270	Thr	Pro
25	Gln	Val	. Asn 275		The	· Ala	тут	Glu 280	ı Met	. As	o Arg	Phe	Phe 285	Ala	Asp	Ile.
•	Gly	Phe 290		: Gl:	: Gly	y Väl	. Val	L Pro	TY	r Ly	s Thi	Arg	Val	Let	ı Pro	Leu
30	<u> </u>		ı Glu	ı Le	ı Me	Th: 310	PT(	o Gl	y Va	l Pr	o Va 31	1 Thr 5	Cys	s Ile	≘ Тут	320
35	<u>A~</u> 2	GL;	y Va	l As	Th 32	r Pro	o Gli	u Va	l Le	и Ма 33	t Ty	r Gly	, Lys	s Gl	y Gl;	y Phe 5
	AS;	, Ly	s Gl	n Pr 34	o Gl O	u Il	e Ly	s Ty	∓ Gl 34	y As 5	ip Gl	y As	9 G1	y Th 35	r Va O	l Asn
-40	Lei	u Al	a Se 35	r Le 5	u Al	a Al	a Le	u Ly 36	rs Va 10	l As	sp Se	r Le	u As: 36	n Th 5	r Va	l Glu
		e As		y Va	l Se	r Hi	s Th	<u>ت</u> S∈ '5	er Il	e L	eu Ly	/s As 38	p Gl O	u Il	e Al	a Lèu
45			u Il	.e												
50	_		3 Q 3 8 9													
			PRT Arai	oido	psis	the	Liana	2								
55	<4 1-5	100> 1 1	Z L	ys G	lu G	ly L	eu Dj	ñe y	la L	Às H	is P 10	ro Va	al Va	1 P	ne I	le Pro 15
60	<b>.</b> 51	<u>'y</u> I	le V	al T	hr G 20	ly G	ly L	eu G	lu L	eu I 25	<b>∵</b> ⊋ G	lu Gl	Ly Ly	ys G	in C; 30	ys Ala



	Asp	Gl	уІ	Leu 35	Phe	<u>A</u>	g L	ys :	Arg	L	eu 40	Trp	Gl	.y (	Gly	Thr	Pł 4	ne I 15	Leu	Cña	Tr	Ð
5	Val	G1 S	u 1	His	Met	Se	r I	'sn	Asp 55	A	sn	Glu	Tì	ır	Gly	en 50	A	g de	Pro	Alā	Gl	.Ÿ
	Ile 65	Ar	g '	Val	Arg	Al	.a. T	7al 70	Ser	G	ly	Leu	. Va	1	Ala 75	Ala	. A:	g g	TYI	Phe	A]	La 30
10	Pro	G.	īλ	Tyr	Phe	Vē 8	al (	LLD	Ala	V	al	Let	ıI	1e 90	Ala	Asn	L	eu.	Ala	His 95	I.	le
1.5				Glu	100	)						10.	,									
15				Gln 115						•	120						_					
20		1	30	Glu					13	<b>-</b>							•					
	143	5		His				150							13-	•						
25				Ala		1	.65						• •	.,0						_		
30				Ty:	18	0													-			
20				Ly:	5						200	,										
35			210						2.2	75												
	22	25		de c				23	O						دے	-						
40				o Gl			245	)						20	•						_	
45				g Al	2	50						2	93						-	_		
40	P:			e G3	75						20	, 0							_			
50	L	eu	Ly 29	s <u>A</u>	la G	ily	۷a	፲ ፻፶	/= A 2	95	Va	Δ. Δ	.sp	Gl	y As	≅⊋ G 3	1u 00	Th	r Va	1 9	ro	Val
	L 3	eu 05	Se	r A	la C	ly	Ty.	= Me	et C	УS	: Al	la I	.ys	Al	.a T:	က္ A 15	rg	G1;	A F?	rs T	in ar	Arg 320
55				m P			32	5						33	30					-	33	Ser
- ند		) <u>=</u> 0	25	c A	la :	Ast. 340	Le	u L	eu (	3lι	ı G	ly !	\ 3 4 5	G	y T	ir G	31n	Se	۳ G: 3:	Ly 3 50	lla.	His
60	) V	7al	. As	sp I	le :	Met	Gl	y A	sn 1	Ph	e A	la i	Leu	. II	le G	14 P	reb	Il	e M	et 2	,_==	Val

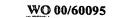


	355 360 365	
	Ala Ala Gly Gly Asn Gly Ser Asp Ile Gly His Asp Gln Val His Ser 370 375	
	Gly Ile Phe Glu Trp 385	
10	<210> 4Q <211> 1986 <212> DNA <213> Saccharomyces cerevisiae	
15	<220> <221> CDS <222> (1)(1983)	
20	atg ggc aca ctg ttt cga aga aat gtc cag aac caa aag agt gdc too atg ggc aca ctg ttt cga aga aat gtc cag aac caa aag agt gdc too Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser 10	18
25	gat gaa aac aat aaa ggg ggt tot gtt cat aac aag cga gag agc aga Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg 20 25	96
30	aac cac att cat cat caa cag gga tta ggc cat aag aga aga agg ggt.  Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly  40  45	144
	att agt ggc agt gca aaa aga aat gag cgt ggc aaa gat ttc gac agg  Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg  50 55	192
35	aaa aga gac ggg aac ggt aga aaa cgt tgg aga gat tcc aga aga ctg Lys Arg Asp Gly Asm Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu 65 70 75	240
40	att ttc att ctt ggt gca ttc tta ggt gta ctt ttg ccg ttt agc ttt Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe 85 90	288
45	ggc gct tat cat gtt cat aat agc gat agc gac ttg ttt gac aac ttt Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe 100 105	336
50	gta aat tit gat toa ott aaa gtg tat tig gat gat tgg aaa gat git Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Tro Lys Asp Val	384
50	ctc cca caa ggt ata agt tog ttt att gat gat att cag gct ggt aac Leu Pro Gin Gly Ile Ser Ser Phe Ile Asp Asp Ile Gin Ala Gly Asn 130	432
53	tac too aca tot tot toa gat gat one agt gas aat tot god got ggt  Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly  145	480
6	The ser are dec dec and cat cet gtt gta	528



### 28/53

					16	5					1	L70						175			
5	atg Met	gtt Val	cct Pro	ggt Gl <sub>3</sub> 180	y Va	c a	itt []e	tct Ser	acg Thr	G.	ga a ly : 85	att Ile	gaa Glu	agc Ser	tg TI		ga 1y .90	gtt Val	at	t.e	576
	gga Gly	gac Asp	gat Asp 195	Gl	g tg	ys i	gat Asp	agt Ser	tct Ser 200	-	cg la	cat His	ttt Phe	cgt Arc	20		ra Saa	Ctg Leu	T:	ed 12	624
10	gga Gly	agt Ser 210	Phe	ta Ty	c a r M	tg et :	ctg Leu	aga Arg 215	aca Thi	a a	tg let	gtt Val	atg Met	gat Asi 220	,,	ta (	gtt Val	cys	T:	rp gg	672
15	ttg Leu 225	aaa Lys	. cat	gt Va	a a l M	tg et	tta Leu 230	gat Asp	Pro	e g	raa Slu	aca Thr	ggt Gly 235	T) C	g ga	g g	cca Pro	ccg Pro		ac sn 40	720
20	ttt Phe	acc	r ct: Le	e co	œ A	ca la !45	gca Ala	cag	gg Gl	c t	tc Phe	gaa Glu 250	tca Ser	ac Th	t ga	at ge	tat Ty <u>r</u>	Phe 255	-	tc le	768
25	gca Ala	GJ7 ggg	, ta , Ty	t to T T: 26	. Q.	itt [le	tgg Trp	aac Asi	aa Ly	s \	gtt Val 265	ttc Phe	cas Glr	laa NAS	t c n L	eu Eg	gga Gly 270		i a	itt :le	815
20	Gly	ta: Ty:	c ga c Gl 27	u P	co i	aat Asn	aaa Lys	ato Met	g ac Th 28	<u> </u>	agt Ser	gct	gcg Ala	g ta a Ty		at sp 85	TTP	age Are	g q g I	et Leu	864
30	gca	ta Ty 29	t tt T Le	a g u A	at sp	cta Leu	gaa Gli	age 1 Are 29	g Al	rc e	gat Asp	agg Arg	j ta j Ty:	c tt r Pi 30	16 1	,pr	aaç Lys	ct Le	a i	aag Lys	912
35	gaa Gl: 30	ı Gl	a at n II	ie G	aa lu	ctg Leu	Pho 31	e Hi	t ca s Gl	aa Lr.	ttg Leu	agi Se:	gg Gl 31	y G.	ea a Lu I	Lys	gt: Val	t tg L Cy	_	tta Leu 320	960
40	at: Il	e Gl	y H	is S	ct	atg Met	G1	t to y Se	t ca r G	ag ln	att	at = I1 33	= F11	t t e T	ac : yr !	::: Phe	at; Me	g aa Ly 33	_	tgg Trp	1008
45	gt Va	c ga 1 Gl	lg g	la (	iu	GŢ7	, br	t ct o Le	1. 1.	Ϋ́	gg: Gl: 34:	, as	t gg n Gl	t g y G	gt ( ly )	egt Arg	G1 35		ಚಿ ಕಿ	gtt Val	1056
	aa As	c ga n Gi	aa c lu H 3	ac a is :	ita [le	gat As;	t to o Se	a to r Pi	ie I	tt 1e 60	AS	t go n Al	a go a Al	a g .a G	<b>∸</b> ⊻	acg Thi 365		t ct u Le	eu eu	Gly	1104
50	gc Al	a P	ca a ro L 70	ys .	acs Siz	gt: Va	t co l Pr	:0 A	ct c la L 75	ta ,eu	at:	- <b>a</b> g e Se	rt gg	Ly G	aa lu 80	atg Met	aa Ly	ag: SA:	at sp	acc Thr	1152
55	a: 11	.e G	ea c ln I	eu :ta	aat Asn	ac Th	g t T Le	eu A	cc a la M	itg (et	ta Ty	t 99 T 63	.y 21	eu 9 95	aa lu	aaq Lys	: 11 : 7:	c t le P	tc he	Ser 400	1200
60	ag Az	ja a :g I	tt g le G	ag lu	aga Arg	gt Va 40	T T	aa a /s M	tg : et I	: T.E.	ca Gl		eg t er Ti	ات اعد د	igt 1y	gg: Gl;	: at / Il		22 20 15	tca Ser	1248







	atg cta cca aag gga gaa gag gtc att tgg ggg gat atg aag tca tct 1296 Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser 420 425	
5	tca gag gat gca ttg aat aac act gac aca tac ggc aat ttc att 1344  Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile  435  440	
10	cga ttt gaa agg aat acg agc gat gct ttc aac aaa aat ttg aca atg 1392 Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met 450 455	
15	aaa gac gcc att aac atg aca tta tcg ata tca cct gaa tgg ctc caa 1440  Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln  470  480	
. 20	aga aga gta cat gag cag tac tog tto ggo tat too aag aat gaa gaa 1488 Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu 495 485	
	gag tta aga aaa aat gag cta cac cac aag cac tgg tcg aat cca atg 1536 Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met 500 505	
25	gaa gta cca ctt cca gaa gct ccc cac atg aaa atc tat tgt ata tac 1584  Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr  515  520  525	
30	ggg gtg aac aac cca act gaa agg gca tat gta tat aag gaa gag gat 1632  Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp  530  530	
35	gac too tot got otg aat ttg acc atc gac tac gaa agc aag caa cot 1680 Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro 545 550 555	
40		
	atg tgt cac aaa tgg gcc cag ggt gct tca ccg tac aac cct gcc gga 1776  Met Cys His Lys Trp Ala Glr Gly Ala Ser Pro Tyr Asn Pro Ala Gly  580  580	
45	att aac git act att gig gaa aig aaa cac cag cca gat cga tit gat 1824 att aac git act att gig gaa aig aaa cac cag cca gat cga tit gat 1824 Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp  595 600 600	
5	and the state of t	
5	geg gag tig aac gat tac atc tig aaa att gca agc ggt aat ggc gat 1920 5 Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp 635 640	)
ć	ctc gtc gag cca cgc caa ttg tct aat ttg agc cag tgg gtt tct cag 1966 Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Try Val Ser Gln 655 645	3



1986 atg ccc ttc cca atg taa Met Pro Phe Pro Met 660 5 <210> 5 & <211> 661 <212> PRT <213> Saccharomyces cerevisiae 10 <400> 5 Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg 20 25 30 Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly 35 40 45 20 Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg 50 55 60 Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu 65 70 75 80 25 Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe 30 Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val 35 Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly 40 145 Lys Gln Leu Lau Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile 45 Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp 50 Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp Deu lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn 225 230 230 240 55 Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile 250 Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile 250 255 270

### **SUBSTITUTE SHEET (RULE 26)**



	Gly	Tyr	Glu 275	Pro	Asn	Lys	Met	Thr 280	Ser	Ala	Ala	Tyr	Asp 285	Trp	Arg	Leu
5	Ala	Тут 290	Leu	qzA	Leu	Gľu	Arg 295	Arg	qaA	Arg	Tyr	Phe 300	Thr	Lys	Leu	Lys
10	Glu 305	Gln	Ile	Glu	Leu	Phe 310	His	Glņ	Leu	Ser	Gly 315	Glu	Lys	Val	Cys	Leu 320
10	Ile	Gly	His	Ser	Met 325	Gly	Ser	Gln	Ile	Ile 330	Phe	Tyr	Phe	Met	Lys 335	Tro
15	Val	Glu	Ala	Glu 340	Gly	Pro	Leu	TYT	Gly 345	Asn	Gly	Gly	Arg	Gly 350	Trp	Val
	Asn	Glu	His 355	Ile	qzA	Ser	Phe	Ile 360	Asn	Ala	Ala	Gly	Thr 365	Leu	Leu	Gly
20	Alā	Pro 370	Lys	Ala	Val	Pro	Ala 375	Leu	Ile	Ser	Gly	Glu 380	Met	Lys	Asp	Thr
25	Ile 385		Leu	Asn	Thr	Leu 390	Ala	Met	Tyr	Gly	Leu 395	Glu	Lys	Phe	Phe	Ser 400
23	Arg	Ile	Glu	Arg	Val 405		Met	Leu	Gln	Thr 410	Tro	Gly	Gly	Ile	Pro 415	Ser
30	Met	Leu	Pro	Lys 420		Glu	Glu	Val	11e 425	Trp	Gly	Asp	Met	Lys 430	Ser	Ser
	Ser	Glu	Asp 435		. Leu	. Asn	Asn	Asn 440		Asī	Thr	Tyr	Gly 445	Asn	Phe	Ile
35	Arg	Phe 450		. Arg	ASD	Thr	Ser 455		Ala	Phe	Asn	Lys 460	Asn	. Leu	Thr	Met
40	Lys 465		Ala	. Ile	e Asn	470		Lev	. Ser	: Ile	8ez 475	Pro	Glu	r wit	Leu	Gln 480
40	ΑΞQ	ı Arç	val	. His	Glu 485		Tyr	Ser	Phe	Gl <sub>3</sub> 490	у Т <u>у</u> т )	Ser	Lys	a Asi	495	ı Glu
45	Glu	Le:	1 A <u>r</u> g	Lys 500		ı Glu	ı Lev	. His	505	Lys	s His	TIP	Ser	Asr 510	Pro	) Met
	Gl	ı Val	1 Pro		ı Pro	o Glu	: Als	Pro . 520	His	s Me	Lys	: Ile	∓ Ty:	r Cys	s Ile	e Tyr
50	Gly	y ₹a: 530		a Ası	n Pro	o Thi	Glu 53:		; Ala	i Tyr	r Val	Ty:		s Glu	ı Glı	ı Asp
	As; 54		r Se:	- Ala	a Le	ı As:		: Thi	- Il	e As	7 55	Glu S	ı Se:	r Lys	s Gl	560
55	Va.	i Ph	e Le	ı Thi	= Gl		y Asi	G Gly	y Thi	- Va 57	l Pro	Lev	· Va	l Ala	8 Hi:	s Ser 5
60	Me	E Cy.	s Hi	58		p Al:	e Gl:	a Gly	y Ala 58	a Se 5	r Pr	ту:	r As	n Pro 590	o Al	s GīĀ

	Ile	Asn	Val 595	Thr	Ile	Val	Glu	Met 600	Lys	His	Gln	Pro	Asp 605	Arg	Phe	Asp
5	Ile	Arg 610	Gly	Gly	Ala	Lys	Ser 615	Ala	Glu	His	Val	Asp 620	Ile	Leu	Gly	Ser
	Ala 625	Glu	Leu	Asn	Asp	Tyr 630	Ile	Leu	Lys	Ile	Ala 635	Ser	Gly	Asn	G1y	Asp 640
10	Leu	Val	Glu	Pro	Arg 645	Gln	Leu	Ser	Asn	Leu 650	Ser	Gln	Trp	Val	Ser 655	Gln
15	Met	Pro	Phe	Pro 660												

```
WO 00/60095
SECUENCE LISTING
<110> Siymne Dr., Sten
<120>
<130>
くよるリン
<141>
 <150>
 <210> 11-
 <211> 1986
 <212> genomic DNA
 <211> Saccharomyces cerevisiae
 <220>
 <221> CDS
 <222> (1) .. (1983)
 <400> 1
 atg ego aca ong tit oga aga aat gto ong and ola ang agt gat tot
                                                                      4 B
 Met Gly Thr Lou Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Sor
                     5
  gat yee aac aat aaa ggg ggt tot gtt cat aac aag oga gag age aga
  Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg
                                                         30
                20
  all the att cat can can eag ggs the ege can as aga aga agg egt
                                                                      144
  Ash His lie His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly
   att agt gge agt gea and aga wat gag egt gge and gat the gad agg
   The ser Gly ser Ala Lys Arg Arn Glu Arg Gly Lys Arp Pho Asp Arg
                                                  60
                             55
        50
   and aga gat ggg and ggt aga and ogt tigg aga gat too aga aga etg
   Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu
                         70
    €5
    att the art out got got the the ggt got the the cog the age the
                                                                        288
```

### SUBSTITUTE SHIELET (RULLE 26)

The Pho The Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe

ggo got tal cat gut cat aat ago gat ago gao tog tit gao aas tut Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe 105

25

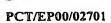
100

9 ℃

## WO 00/60095



Sta ast tit gat toa cit and Sig tat tig gat gat tig and gat Sit Val Asn Phe Asp Sor Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val	364
cuc cca can ggt ata agt tog the act gat gat att cag get ggt aac Lou Pro Gln Gly Ile Ser Sor Pre Ile Asp Asp Ile Gln Ala Gly Asn 130	432
tac tee acc tet tee tta gat gat etc agt gas aat ttt ges git ggt Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly 145 150 150	130
DAR CAR OTO DEA OGT GAT DAT ART DEC GAS GOO DAR CAT OUT GLE GLE Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Alt Lys Ris Pro Val Val 165	520
acg got cot ggt gte att tot acg ggt att gas ago tgg ggs got tot Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile 180	<b>37</b> 6
est sac est ese est agr ret eng cat tot ogt ass con obj tes Gly Asp Asp Glu Cys Asp Ser Ser Ala Ris Phe Arg Lys Arg Leu Trp 195 200 205	624
ggs agt ttt tac atg ctg aga aca atg gtt atg gat aaa gtt tgt tgg Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp 210 215 220	672
tig aas cat gos atg tts gat cot gas aca ggt otg gad osa cog aac Lou Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn 235 230 235	720
tet acg cta cgt gca gca cag ggt tte gas tes act gat tat tte ate The Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile 245	768
ges agg the tag att tag are mad off the cas art sta agg are att and cas art sta agg are att and cas art sta agg are att and cas are all all als all agg are are agg are are all all all agg are are agg are are all agg are are agg. Als are are agg. Als ar	: 515 e
ggo tat gas due aat das aug aeg agt got gog tat gat igg agg of Gly Tym Glu Pro Asn Lys Met Thm Ser Ala Ala Tym Asp Tmp Arg Le 285	E 854



GCA tat tta gat cca gaa aga cgc gat agg tac ttt acg ang cta ang 912  Ala Tyr Leu Asp Leu Clu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys  290 295	
gaa caa atc gaa ctg ttt cat caa ttg agt ggt gaa aaa gtt tgt tta 960 Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu 310 315 320	
att gga cat tot atg ggt tot dag att atd tit tad tit atg all tigg 1008  Ile Gly Kis Ser Met Gly Ser Glm Ile Ile Phe Tyr Phe Met Lys Trp  330  335	
gtc gag gct gaa ggc cot ctt tac ggt aat ggt ggt cgt ggc tgg gtt 1050 Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val  345	
and gas can are gat the the ath ast got got ggg and the ctg ggc 110 Ash Glu Mis Ile Asp Ser Phe Ile Ash Ale Ale Gly The Leu Deu Gly 355	4
get een mag gen git een get ein att agt ggi gan mig man gat mee lis Alm Pro Lys Alm Val Pro Alm Leu Ile Ser Gly Glu Met Lys Asp Thr 370 375	2
att cas tts sar acg tts god atg tat cgc ttg gas aag ttc ttc tca 120  Ile Glm Leu Asm Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser  400  385	00
age att gag age gts and atg tte cas ang tgg ggt ggt ate cos ton 12 Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Scr 415	48
Atg cta cos and ega gas gas ett tes egg gat atg ang ten tet 12  Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser  430	296
toa gag gat goa tog ant aac and not gad aca tad ggd aat tto att 1 ser Glu Asp Ala Deu Asm Asm Asm Thr Asp Thr Tyr Gly Asm Phe Ile 445	344
ogs tit gas agg ast acg ago gat got too aso has ast tog aca atg — I Arg Phe Glu Arg Ash Thr Ser Asp Ala Phe Ash Lys Ash Leu Thr Met 450 450	392
Lys Asp Ala Ile Ash Met Thr Leu Ser Ile Ser pro Glu Trp Leu Gln 475	1440

2	- 1	"	2
		•	•
•	"	•	_

aga ago gt								Asn G	Jr G		i
	4	185			490			4	. 9 5		
gag tta ag Glu Lou Ar	a eda a a eva a	azt gag Asn Glu	cta cac Leu His	cac Hiz	ass o	ac tg	b sei	aat o	eca a:	:g 1536	5
	500			205	-			510			
gsa gta co	a ctt	cca gaa	get ced	CAC	atg a	aaa at	ic tat	tgt a	ETA T	ac 158-	4
Glu Val Pr El		sko ern	520 520		WGC 1	nya ma	525	<b>L</b> , <b>D</b> .		. ~	
599 919 A2	25 8.20	ccz act	gaa ag	e gez	tat	gta ta	at aag	G55 (	525 E	at 163	2
Gly Val As	n Asn	Pro Thr	Glu Arg	g Ala	īār,		yr Lys 40	GIU	GLU A	5 <b>D</b>	
gat tet t:	et get	cts aat	ttg ac	c atc	ÇEC	tac g	aa agc	aag	caa c	c: 168	C
Asp Ser S: 545	er Ale	Leu Asn	Leu Th	r ile		ESS	16 241	دولد	5	50	
gta tto c	tc acc	2 <i>4</i> 2 882	esc ee	a acc	gtt	ceg c	te gts	gcg	cat t	ca 172	8.8
Val Phe L	ou Thr	Glu Gly	Asp Gl	y Thr	V21 570	Pro L	en Avt	ALE	575	: <b>8</b>	
ate tet c	ac aaa	Cgg gcc	cag gg	t got	. cca	ces t	ac zac	cct	gcc g	ga 17:	7€
Met Cys R	is Lys 580	Trp Ala	Glæ Gl	ese Ese		Pro 1	ryr Asi	590	ا علم	±17.	
att aac g	tt act	att gt	gaa at	tg ata	cas	cag (	cca ga:	cga	ttt	gat 167	24
Ile Asn V	Tal Thr 195	Ile Va		at Lys oo	: His	Gln I	Pro Asi		Phe .	Asp	
ata ogt s	gc gga	. gca aa	a ago g	cc gaa	z cac	gta 9	gac at	c ctc	55c	age 18	72
Ils Arg (	sly Gly	· Ala Ly	s Ser A 615	la Gli	ı His	Val 2	Asp Il 620	e Leu	Gly	Ser	
608 898	ttg zzc	gat ta	c atc t	tg za	a att	gea	ago gg	t aat	gge	çat 19	20
Ala Glu : 625	leu Asn	Asp Ty 63		eu ly	s Ile	635	ser Gl	y Ass	. Giy	640	
ato ges	gag cca	s cac ca	.z ttg t	ct ca	.t ===	gago	cag tg	g gtt	tet	cag 15	962
Leu Val	Glu Pr	0 Arg Gl 645	.r Leu S	er As	n Dev	: Se=	Gln T	p Val	Ser 655	Gjm	

# SUBSTITUTE SHEET (RULE 26)

/53

1986

atg ccc ttc cca atg taa Met Pro Phe Pro Met 660

<210> 2 ls
<211> 651
<212> PRT
<213> Sacchiromyces cerevisiae
<213> \$2000000000000000000000000000000000000
<400> 2
Met Gly Thr Lou Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser
Met Gly The Lett Fite Alg 712 10
1 5 San Lys Arg Glu Ser Arg
1 5 Asp Glu Asp Asp Clu Ser Arg Asp Glu Asp Asp Glu Asp Asp Glu Asp
20 Cly Fig Lys Arg Arg Gly
20 23 Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly
25 40
15 40 Lic Ser Gly Ser Ale Lys Arg Ash Glu Arg Cly Lys Asp Phe Asp Arg 60
50 55
50 E5 Lys Arg Asp Gly Ash Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu 75 80
75
65 70  The Pas The Leu Gly Ala Phe Leu Gly Val Leu Lau Pro Pas Ser Phe  10 90 95
The Pas the Lett City Ale to
S5  Gly Ala Tyr His Val His Asn Ser Asp Scr Asp Leu Phe Asp Asn Phe  110
Gly Ala Tyr His Var His Ash 105
100 The Let Asp Asp Trp Lys Asp Val
Val Ash Pho Asp Ser Lou Lys Val Tyr Leu Asp Asp Trp Lys Asp Val
115 120 Ash Tie Gin Ala Gly Ash
115  Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Ass  140
130 135 an and ala Val Gly
135  136  137  Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly  150  150
Tyr ser int set at 255 150
145 150 150 Lys Glu Ala Lys His Pro Val Val Lys Glu Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val
170
165  Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile  190
Met Vai pro Gly vai 120 1 185
Gly Asp Asp Glu Cys Asp Ser Ser Ala Mie Phe Arg Lys Arg Leu Trp
Gly Asp Asp Glu Cys Asp 322 200 205
195 The Mer Val Met Asp Lys Val Cys Tip
195 200 Cys Trp Gly Ser Pne Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp 220
210 215 ASD Pro Pro ASD
215 215 Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn 240
230 230 The len Tyr phe Ile
225 230 230 Phe Glu Ser Thr Asp Tyr Phe Ile Phe Thr Leu Arg Alz Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile
250
245  Ala Gly Tyr Trp Ile Trp Asn Lys Val Pne Gln Asn Leu Gly Val Ile 253 270
A12 Gly 171 115 115 265 270
260 225 Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu
Gly Tyr Glu Pro Ash Dys Med 122 0 285
275 200 let Lys let Lys Let Lys
275 280 Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys 300
295  Glu Gln The Glu Leu Phe Mia Gln Deu Ser Gly Glu Lys Val Cys Deu 320
Giv Gla The Glu Leu Phe His Gla Leu Ser Gly Garage 320
710
305

The MCE Lys Tra
Ile Gly His Ser Met Gly Ser Gln Ile Ile Pho Tyr Phe Mot Lys Trp
Val Glu Ale Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val
340
Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Lou Gly  355
NSR GIU AIS 222 1125 - 360
Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Mat Lys Amp The
Ala Pro Lys Ala val 120 130
370 The Tou his Met Tyr Gly Leu Glu Lys Pho Phe Ser
The Gln Leu Ash The Leu All 195 400
Arg Ile Glu Arg Val Lys Met Leu Glr. Thr Trp Gly Gly Ile Pro Ser
Arg Ile Glu Arg Val Lys Met Hed Catt
405  Het Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Mct Lys Ser Ser  430
Het Leu Pro Lys Gly Glu Glu Vai 110 129 430
420 The Ben The Tyr Gly Asn Phe Ile
420  Ser Glu Asp Ala Leu Ash Ash Ash Thr Asp Thr Tyr Gly Ash Phe Ile  445
435 436 Arg Pho Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met 460
Arg Pho Glu Arg Asn Thr Ser Asp Arg Pho 1111 1
450 A55
455  455  Lys Asp Ala Ile Asn Met Thr Leu Sor Ile Ser Pro Glu Trp Leu Gln 475  475
465 470 Ser Lys Asn Glu Glu
465 470 Arg Arg Val His Glu Gln Tyr ser Phe Gly Tyr Ser Lys Asn Glu Glu 490 495
485 490 Ser Ash Pro Not
485  Glu Lou Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Mct  505  510
500 505 The Tyr Cvs Ile Tyr
Glu Val Pro Lou Pro Glu Ala Pro Hic Met Lys Ile Tyr Cys Ile Tyr
Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp
540 535
535 Asp Ser Ser Ala Leu Ash Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro 555
555 555 550 555
545 550 541 Pro Leu Val Ala His Ser Val Pro Leu Val Ala His Ser Val Pro Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser
565
565  Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly  590
885 S85
580  See See See See See See See See See Se
11c Asm Val Ini 134 Val 600
595 600  Ile Arg Gly Gly Ala Lyr Ser Ala Glu His Val Asp Ile Leu Gly Ser  620
ile Arg Gly Gly All Dyl 500 620
fin 615 Ala Glu Leu Ash Asp Tyr Ile Leu Lys Ile Ala Ser Gly Ash 640
Ala Glu Leu Ash Asp Tyr 112 Dec 275 635
630 630 Leu Ser Glm Trp Val Ser Glm
625 630 Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln G55
645
Met Pro Phe Pro Met
650

PCT/EP00/02701

WO 00/60095

<210> 3 br <211> 2312 <212> genomic DNA <213> Schirosaccharomyces.pombe <400> 3	
NTOGOGTOTT CCAAGAAGAG CAAAACTCAT AAGAAAAAA AAGAAGTCAA	50
ATCTCCTATC GACTIACOAA ATTCAAAGAA ACCAACTCCC GCTTTGAGTG	100
AGCARCOTTO AGOGTOOGAA ACACARTOTG TTTOARATAR ATCARGARAR.	150
TCTAAATTTG GAAAAAGATT GAATTTTATA TTGGGCGCTA TTTTGGGAAT	200
ATGEGGTGET TTTTTTTEG CTGTTGGAGA CGACAATGET GTTTTCGACC	250
CTGCTACGTT AGATAAATTT GGGAATATGC TAGGCTCTTC AGACTTGTTT	300
GATCACATTA AAGGATATTT ATCTTATAT GTGTTTAAGG ATGCACCTTT	350
TACTACGGAC AAGCCTTCGC AGTCTCCTAG CGGAAATGAA GTTCAAGTTG	400
GICTTGATAT GTACAATGAG GGATATCGAA GTGACCATCC TGTTATTATG	450
GTTCCTGGTG TTATCAGCTC AGGATTAGAA AGTTGGTCGT TTAATAATTG	500
CTCGATTCCT TACTTTAGGA AACGTCTTTG GGGTAGCTGG TCTATGCTGA	550
AGGCAATGTT CCTTGACAAG CAATGCTGGC TTGAACATTT AATGCTTGAT	€00
ALAXARACCO GCTTGGATCO GRAGGGAATT AAGCTGCGAG CAGCTCAGGG	650
GTTTGARGON GCTGRTTTTT TTRTCACGGG CTRTTGGRTT TGGRGTRARG	700
TARTTGARAR CCTTGCTGCR RTTGGTTRTG RGCCTRRTAR CRTGTTRAGT	750
CCTTCTTACG ATTGGCGGTT ATCATATGCA AATTTAGAGG AACGTGATAA	200
ATATTITICA ANGITAAANA IGITCATIGA GTACAGCANC ATIGIACATA	250
AGAAAAAGGT AGTGTTGATT TCTCACTCCA TGGGTTCACA GGTTACGTAC	900
TATTTTTTA AGTGGGTTGA AGULGAGGGC TACGGAAATG GTGGACCAAC	950
ITGGGTTAAT GATCATATTG AAGCATTTAT ARATGTGAGT CTCGATGGTT	1000
GTTTGACTAC GTTTCTAACT TTTGAATAGA TATCGGGATC TTTGATTGGA	1050
GCACCCARAR CRGTGGCAGC GCTTTTRTCG GGTGARATGA AAGATACAGG	1100
TATIGTAATT ACATTAAACA TGTTAATATT TAATTTTTGC TAACCGTTTT	1150
ANGETCARIT GAATCAGITI TEGGTETATG GGTAAGCAAT AAATTGITGA	1200
GATITUTIAC TAATTTACTG TITAGTTTGG AAAAATTTT TICCCCTTCT	3250
GAGGTATATT CARARTACA ARTGTGCTCT ACTTTTTCTA ACTTTTARTA	1300
GAGAGOCATO ATGGTTOGOA CTATGGGAGG AGTTAGTTOT ATGCTTOCTA	1350
ARGRAGICA TETTETATES EGARATECER ETTEGETARE ARMATETEC	1400



GTTAATTTT TIATTAATAT TTAGGCTCCA GATGATCTTA ATCAAACAAA	1430
TTTTCCNAT GGTGCAATTA TTCGATATAG AGAAGACATT GATAAGGACC	1500
ACGATGARTT TGACATAGAT GATGCATTAC AATTITIAAA AAATGTTACA	1550
SATGACGATT TTAAAGTCAT GCTAGCGAAA AATTATTCCC ACGGTCTTGC	1600
TTGGACTGAA AANGAAGTGT TAAAAAATAA CGAAATGCCG TCTAAATGGA	3,650
TANATOCGOT AGNAGTAAGN ACATTAAAGT TACTAAATTA TACTAACCCA	1700
ANTAGACING TOTTCOTTAT GCTCCTGATA TGANAATTTA TIGCGTTCAC	1750
GGGGTCGGAA AACCAACTGA GAGAGGTTAT TATTATACTA ATARTCCTGA	1500
GGGGCAACCT GTCATTGATT CCTCGGTTAA TGATGGAACA AAAGTTGAAA	1850
ATGTGAGAGA ATTTATGTTT CAAACATTGT ATTAACTGTT TTATTAGGGT	1900
ATTGTTATGG ATGATGGTGA TGGAACTTTA CCAATATTAG CCCTTGGTTT	1950
GGTGTGCAAT AAAGTTTGGC AAACAAAAAG GTTTAATCCT GCTAATACAA	2000
GTATCACARA TTATGRARIC RAGCATGRAC CTGCTGCGTT TGATCTGAGA	2050
GENERALE CAROCATORS GENERALE GETEGGERA ACREGICGAT ATACTIGGRE ATTENGAGET	2100
GGNGGACCTC GCTCGGCAGA ACACGTCOTO	2150
AARTGTATET TCATTTACC TTACACCTTT  AAGGANATTA TTTTAAAAGT TTCATCAGGC CATGGTGACT CGGTACCAAA	2200
AAGGANATTA TITTAAAAGI TICATCAGG COOLAGATTIGT AGATTGCAAT CCGTTATATA TCAGATATCC AGTACGGACA TAAGTITTGT AGATTGCAAT	2250
CCGTTATATA TCAGATATCE AGIACGGACA INCOPPA TAACTAACTA ACCGAACAGG GAAATAATAA ATGAGATAAA TCICGATAAA	2300
TARCTARCTA RECEARCAGE GARALARIAN ALCOHOLING	2312
CCTAGAAATT AA	

<210> 4 b <211> 3685 <212> genomic DNA <213> Arabidopsis thaliana <400> 4

TOTAL CONCERNATION CONCECCED	30
TGCCCCTTA TTCATCGGAA AAAGCCGACG GAGAAACCAT CGACGCCGCC	100
TGCCCCTTA TTCATCGGAA AAADOO	150
CTTCCARATC CCACCATARG ARATCGARCG GAGGAGGGAR GTGGTCGTGC	200
ATCGATTCTT GTTGTTGGTT CATTGGGTGT GTGTGTAA CCTGGTGGTT	250
ATCGATTCTT GTTGTTGGT: CATTOCCTGC GAGCTTCCCT CAGTATGTAA TCTTCTCTTC CTTTACAACG CAATGCCTGC GAGCTTCCCT TAAGCTCAAA	300
COGREGORAT CACEGOTOCT TEGCTERCE COCCEGOTOT TARGETCARA	350
CAGAGGAAT CACAGGGICCT TO CAGAGATTAT CAAAGAAGGTC TTAAGGCGAA ACATCCTGTT GTCTTCATTC CTGGGATTAT	400
AAAGAAGGTC TTAAGGCUAR ACATTOCGCT GATGGTTTAT CACCGGTGGG CTCGAGCTTT GGGAAGGCAA ACARTGCGCT GATGGTTTAT	450
GGAACTTTTG GIGARGIOTH	500
TOTAL CONTRACT GGGAIII	550
TECTICALA TEATION	600
TECTITITE 1111	650
CTGTATGTGT KAGGT	700
CTGTTIGATI GILARCI	750
TOTAL STORTSTAT TOCCOMME	€00
THE PROPERTY OF THE PROPERTY O	250
CGTATARTER AGIRGICA	900
TECATOTET GOALS	950
TTATTCAACT ALGIRE	1000
THE PROPERTY OF THE PARTY OF TH	1050
- CARCETTATT CALKICION	1100
COTTERACT COLUMN	
TREEGESTE CALCULATION	1150
	1200
	1250
ARTGITATIG TIGATGGTAR CAGIGOTATO TIGGGTGGAR CACATGICAC TCARTATOT CITTGAITCR GGCCTCTATG TIGGGTGGAR CACATGICAC	1300
TCANTATOT CITTONITCA GGCCTCINIO TOTALIAGAGI TCGAGCTGIA:	1350
TIGACARIGA ARCIGGGING GATCONGCIG GIALES	

THE	1400.
TCAGGACTCG TGGCTGCTGA CTACTTTGCT CCTGGCTACT TTGTCTGGGC	1450
AGTGCTGATT GCTAACCTTG CACATATTGG ATATGAAGAG AAAAATATGT	1500
ACATGGCTGC ATATGACTGG CGGCTTTCGT TTCAGAACAC AGAGGTTCTT	1550
TTCTCATCGT TCTTTCTATT ATTCTGTTCC ATGTTACGTT TCTTTCTTCA	2600
TILETTARGE CTIANATATE TITCATETTE ARTTAATAGE TACGIGALCA	1650
CACTOTTAGO CÓTATGAAAA GTAATATAGA GTTGATGGTT TCTACCAACG	·
GTGGARRAR AGCAGTTATA GTTCCGCATT CCATGGGGGT CTTGTATTT	1700
CTACATTITA TGAAGTGGGT TGAGGCACCA GCTCCTCTGG GTGGCGGGG	1750
FEGGCCAGAT TGGTGTGCAA AGTATATTAA GGCGGTGATG AACATTGGTG	1900
CACCATITET TESTETTECA AAAGCTETTE CAGGGCTTTT CTCTGCTGAA	1850
CONSCRITE TECRETTEC CAGGTATTEA ATATCTECTT ATACTITIES	1900
TGATCAGAAC CTTGGCTCTG GAACTCAAAG TTATTCTACT ARATATCAAT	1950
TOTALTANCA TIGOTATATI ATCGCTGCAA CTGACATTGG TIGATTATIT	2000
TETARTANCA TIGETATORE ACTETETES GATTAGACAA ATGATGAATT TTGCTGCTTA TGTAACTGAA ACTETCTTGA GATTAGACAA ATGATGAATT	2050
GATAATTOTT ACGCATTGCT CTGTGATGAC CAGTITCTTA GCTTCGACGA	2100
GATAATTCTT ACGCATIGCT CTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTO	2150
TAACATTIGT CATACTGICI IIIGGADDD  GCGCIGGAGC TTCCATGCTT GCATTCTTTA CCAATTAGCG TTATTCIGCT	2200
GCGCTGGAGC TTCCATGCTT GCATCTATG GTCTTTATT TCTTCTTAAT TCTTTCAATT TTCTTGTATA TGCATCTATG GTCTTTTATT TCTTCTTAAT	2250
TOTTTCARIT TICTIGIATA IGCALCIATO	2300
TARAGACTOG TTGGATTAGT TGCTCTATTA GTCACTTGGT TCCTTAATAT	2350
AGAACTITAC TITCTICGAA NATIGCAGAG CGATTGCCCC AGGATTCTTA	2400
GACACCGATA TATTTÄGACT TCAGACCTTG CAGCATGTAA TGAGAATGAC	2450
ACCCACATGG GACTCAACAA TGTCTATGTT ACCGAAGGGA GGTGACACGA	2500
TATEGEGEGE GETTEATIEG TEACEGEAGA AAGGECACAE ETGTTGTGGG	2550
ANALGONAL AGARCARCGA ARCTTGTGGT GARGCAGGTG ARARCGGAGT	2600
TTCCAAGAA AGTCCTGTTA ACTATGGAAG GATGATATCT TTTGGGAAAG	
North Color Goorgeoock Tetgagatta Ataatattga Titlegagia	
MODICATED ARTCATARTA RACCTTGTAC ATTTTGTGAT TGIRLGATOR	
TO THE TAXABLE GIGALGGIG CIGICALAGG TOAGAGIALD	
COLLEGE COTOTOGTGA COTOTOGACA GAGTACCATG ACAIGGGTAC	
TGCTGGGATC AAAGCTATCG CTGAGTATAA GGTCTACACT GCTGGTGAAG	2650
TOUTGSGAIL ACADELATE	

# SUBSTITUTE SHEET (RULE 26)



TATAGATCT ACTACATTAT GTTGCTCCTA AGATGATGGC GCGTGGTGCC	2900
CTCATTICT CITATGGAAT TGCTGATGAT TTGGATGACA CCAAGTATCA	2950
AGATECEAAA TACTGGTEAA ATECGTTAGA GAEAAAGTAA GTGATTTETT	3000
AGATCCCAAR TACTGGTCAR REGION REGION REPORTED TO THE STATE OF THE STATE	3050
GATTCCAACT GTATCCITCG TCCTCAAA GCTTACAAAG CTGTTTCTGA	3100
GGTCTTGTTG GATATGGTTT TCAGGTAATA TTGAGGTGCT AAAGTTGATA GCCTTTCTCA AAAAGGCTTG CTCAGTAATA TTGAGGTGCT AAAGTTGATA	3150
GOOTTTOTOA ARANGGETTG CICAGITATOR OF CATGIGACTO TIGOTTATAA ATCCTCCGTT TGGTTTGTTC TGCTTTTCA	3200
CATGIGACTC TIGCTTATAR AIGCTECT ACTCATTATA CGGAGTGGGG GATTACCGAR TGCTCCTGAG ATGGARATCT ACTCATTATA CGGAGTGGGG	3250
GATTACCAACGG RACGAGCATA CGTATACAAG CTTAACCAGT CTCCCGACAG	3300
ATACCAACGG AACGAGCATA CGTATACATO OF TACCAACGA GACGAAGATA TTGCATCCCC TTTCAGATAT TCACTTCTGC TCACGAGGAG GACGAAGATA	3350
TIGCATCCCC TITCAGATAT TEACHTCIGE ATGCGGATGA AACAGTACCC GCTGTCTGAA AGCAGGAGTT TACAATGTGG ATGCGGATGA AACAGTACCC	3400
GCTGTCTGAA AGCAGGAGTT TACAAIGIGG AIGAGACAAG GTCCTAAGTG CCGGGTACAT GTGTGCAAAA GCGTGGCGTG	3450
GTCCTAAGTG CCGGGTACAT GTGTGCAACAT COO	3500
ATTCAACCOT TOOGGAATCA AGACTTATAT AAGAGAATAC AATCACTOTO	3550
CGCCGGCTAR CCTGTTGGAR GGGCGCGGGR CGCRGAGTGG TGCCCATGTT	3600
GATATCATGG GARACTTTGC TTTGATCGAR GATATCATGA GGGTTGCCGC	3550
CGGAGGTAAC GGGTCTGATA TAGGACATGA CCAGGTCCAC TCTGGCATAT	368
TIGANTEGIC EGAGCETATI GACCIGAAGC TETGA	



210>	402	1														
<212; <213;	AT:	zbido	psis	cha	lian	Lā.										
<220:			•													
<221:	CD	s														
		20)	. (402	2)												
< 400:																
		CICT														60
CTGG	ACGA	GATT	TGAC	Ð <i>⊈4</i> _4	TCC	GTAT	AGCT	TAAC	CTGG	T TT	AATT	TCAA	STGA	CAGA	T	119
		CTT		~ n ~	ccc	. A A A	AAG	CCG	ACG	CAG	AAA	CCA	TCG .	ACG	CCG	167
ATG	220	Leu	ALL I	uri s	Ara	Lvs	Lys	Pro	Thr	Glu	Ly≤	Pro	Scr	Thr	Pro	
		GAA Glu							Car	TCG	CAA	AAG	AAA	CCA	CAC	215
Pro	Scr	Glu	Glu	VZI	Var	HIE	Map				•					
GAA	TCT	TCC Ser	AAA	TCC	CAC	CAT	AAG Lvs	727 755	TCG Ser	AAC Asn	CJλ GGỳ	GGA Gly	GGG	AAG Lys	TGG	263
GIG	Ser	Ser	₽À=	3=1	***=		-,-									
TCG	TGC	ATC Ilc	GAT	TCT Ser	TGT Cys	TGT	TGG TEP	TTC Phe	ATI Ilc	GJÀ GGG	TGT Cys	GTG Val	CAS	Val	Thr	. 311
									~~~	BTG	CCT	GCG	AGC	TTC	CCT	259
Irp	Trọ	Phe	Leu	Leu	Phe	Leu	TAT	17511			-				Pro	
		GTA				227		CNO	TCC	TTT	GCC	TIA	CCC	G		402

<210 <211 <212 <213 <220	> 64 > cD > 2=	3	Уs													
<221	> (1 > (2	s ),,(	402)											·		
CGG AIS	GAG Glu	nna. Lys	ATA Ile	GCT Ala 5	GCT Ala	TTG Leu	AAG Lys	GGG Gly	GGT Gly 10	GTT Val	TAC Ty:	TTA Leu	GCC Ala	GAT Asp 15	GGT	48
GAT qa.A	GAA. Glu	ACT Thr	GTT Val 20	CCA Pro	GTT Val	CII	AGT Ser	GCG Ala 25	GGC Gly	TAC Tyr	ATG Mei	TGT Cys	GCG Ala 30	aaa Lys	GGA Gly	96
TGG Trp	yrg	35 Gly GGC	AAA Lys	ACT Thr	CGT Arg	TTC Phe	AGC Ser 40	CCT PTG	SCC	GGC	AGC Ser	AAG Lys 45	ACT The	TAC	GTG Val	144
AGA	GAA Glu 50	TAC Tyr	AGC Ser	CAT His	TCG Ser	CCA Pro SS	bro CCC	TCT Ser	ACT	CTC Leu	CTG Leu 60	GAA Glu	GGC	AGG Arg	GSC	192
NCC Thy 65	CAG	Ser. VCC	GGT	GCA Ala	CAT His 70	Val	GAT Asp	ATA	ATG Mec	GGG Gly 75	Asn	TTT	GCT Alz	CTA Leu	ATT Ilc 50	240
GAG	GAC Asp	GTC Val	ATC Ile	AGA PTG ES	Ile	GCT Ala	GCT Ala	GGG Gly	GCA Ala 90	Thr	Gly	GAG Glu	GAA. Glu	ATT Ile 95	GGT Gly	258
GGC Gly	GA1 Asg	CAG Gln	GTT Val 100	Tyr	TCA Set	. GAT Asp	ATA Ile	TTC Pne 105	Lys	TGG Trp	TCA Sez	. GAG	110		AAA Lys	336
		Lev 115	1	. cc	TATO	ggap	GII		DAA	TGCC	GACC	CG I	TIAT	TGCG	TTCC	391
نعم	AGTG	TCCT	GCC:	rgagt	GC A	ACT:	TGG	kt ti	TGC	LAAT	, TAC	rigi/	·ATT	TŢTC	CACGC	449
J.T.	CATT	CGTC	CCT	TGT	CAA A	LTTT:	LCAT:	rt G	ACAGO	SACG	ב פאו	RTGC(	JATA	CGAT	IGTIG	507
TP.	ccec	INII	TTC	AGCA:	rte :	TATA:	TAA	AC TO	CATE	LGGI	s TA	AGTT	JCAT	TTG	CCAGC	565
TG	AAAT	TGTG	TAG	rcst	FIT .	CTTT	ACGA:	et e	LATA	אכאאי	e TG	GCGG.	AGCN	GŢG	CCCA	523
																643





<210> 7 \( \begin{align\*} & 7 \( \begin{align\*} & 211 > & 115 \\ <212 > & PRT \\ <211 > & Zee mays \\ <400 > & 7 \end{align\*}

Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly
1 10 15

Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly

Trp Arg Cly Lys Thr Arg Phe ser Pro Ala Gly ser Lys Thr Tyr Val

Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Lou Leu Glu Gly Arg Gly 50 55 60

Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Pho Ala Leu Ile 65 70 75 80

Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly so so so

Gly Asp Gln Val Tyr Ser Asp Ile Dhe Lys Trp Ser Glu Lys Ile Lys 100 105

Leu Lys Leu 115 

(	GGTGGCGAAG	ACGANGGCGG	AAGTTGGAGG	CTARCGAGAA	TGACNCTCGG	.50
	AGATGGATCT	ACCCTCTAGA	GACACGACTA	CONTIGCACO	CAGCCTCAAG	ioo
	GTHTACHGTT	THTATGGGTA	GGAAGCCGAC	GGAGCGÁGCC	TACATCTATC	150
	TGGCGCCCGA.	TCCCGGGACG	ACAACGCATC	TTTAGATGAC	GATCGATACG	200
	ACTTTGACTN	AGGGGCACAT	TGACCACGGT	GTGATTTTGG	GCGAAGGCGA	250
	TGGCACAGTG	AACCTTATGA	GTTTGGGGTA	CCTGTGCAAT	ANGGGGTGGA	300
	aaatgaacag	ATACASTCCT	GCGGGCTCAA	ARRITARCOGT	GGTCGAGATG	350
	CCGCATGAAC	CAGAACGGTT	CAATCCCAGA	. GGAGGGCCGA	ATACGGCGGA	500
	CTTAARTATG	TAGAAAAGGT	TGAAATTTAT	GAAGAGTAAT	TAARTACGGC	550
	ACATAGGTTA	. CTCAATAGTA	. TGACTAATTA	TTAAAAAAA,	TITTTCIAR	600
	********	AAAAA .				616

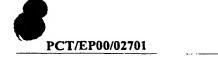
<400> 9



<210> 9 \( \cdot \)
<211> 1562
<212> genomic DNA
<213> Arabidopsis thaliana

atgraraaa :	INICTICACA 1	TATTCGGTA	GTCATAGCGA	TACTCGTTGT.	50
GCTGACGNTG	ACCTCGATGT (	TCAAGCTGT	GGGTAGCAAC	GTGTACCCTT	100
TGATTCTGGT	TCCAGGAAAC (	GGAGGTAACC	AGCTAGAGGT	ACGGCTGGAC	150
AGAGAATACA	AGCCAAGTAG '	IGTCTGGTGT	AGCAGCTGGT	TATATCCGAT	200
TCATAAGAAG	AGTGGTGGAT (	GGTTTAGGCT	ATGGTTCGAT	GCAGCAGTGT	250
	CTTCACCAGG '				300
	TGGATGATTA				350
	GGTTCGACCA				100
				ATAATGAACA	450 
	AAATTTGGGG				500
	AATGTGAGTG				550
				ANTANTGTTG	600
				TTAATAATAA	650
				TGGTGlllGC	700
				CTAGGAGCTC	750
				GTCCCGTGTA	800
				AAACTAGCAG	850
				CTAGGAGGAC	.900
				GCGCCGCAAG	950
TACATCARAC	ACTITGTTGC	ACTCGCTGCC	CCATGGGGT	GGACGATCTC	1000
TÇAGATGAAG	ACATTTGCTT	CIGGCAACAC	ACTCGGTGT(	C CCTTTAGTTA	1050
				TAACCAATGG	1100
				C CCCTTGTCGT	
AACTCCCCAC	GTTANCTAC	a cagettace	QATGGATCG	G TTTTTTGCAG	1200
ACATTGGATT	CTCACAAGG	errerectr	T ACAAGACAA	G AGTGTTGCCT	1250

50/53 ---



	TTANCAGAGG	AGCTGATGAC	TCCGGGAGTG	CCAGTCACTT	GCATATATGG	1300
	GAGAGGAGTT	GATACACCGG	AGGTTTTGAT	GTATGGAAAA	GGAGGATTCG	1350
	ATAAGCAACC	AGAGATTAAG	TATGGAGATG	GAGATGGGAC	GGTTAATTTG	1400
	GCGAGCTTAG	CAGCTTTGAA	AGTCGATAGC	TTGAACACCG	TAGAGATTGA	1450
•	TGGAGTTTCG	CATACATCTA	TACTTAAAGA	CGAGATCGCA	CTTAAAGAGA	1500
	TTATGAAGCA	GATITCAATT	ATTAATTATG	AATTAGCCAA	TGTTAATGCC	1550
	GTCAATGAAT	GA				
						1552

2500

2700 2750

<210> 10 b



<211> 3995 <212> genomic DNA <213> Arabidopsis thaliana <400> 10 ATGGGAGGGA ATTCGAAATC AGTAACGGCT TCCTTCACCG TCATCGCCGT 50 TTTTTCTTG ATTTGCGGTG GCCGAACTGC GGTGGAGGAT GAGACCGAGT 100 TICACGGCGA CTACICGAAG CTATCGGSTA TAATCATTGG CCCAFTTGGS 150 TCGACGCAGC TACGAGCGTG GTCGATCCTT GACTGTCCAT ACACTCCGTT 200 CGACTTCAAT CCGCTCGACC TCGTATGGCT AGACACCACT AAGGTCCGTG ATCITCATTT CCTTCGCTCC TTATTCTGTC GGTCGAGTCA CTTGTTGATG 300 350 TCGTTCATTA GTCARCAGTG ACGCTTCTGA ATCTGAGTTT AGAGTCATAT 400 AAAACAGCYO ACTCGGCGAG TCTTTCCCAT CGCTTTTGGT TCGCTAAATG 450 TAGCGCANTG ANTGTGTANT TAGTCTGCGC TTTTTATTCA ACTAGNICIG 500 CAAGTTTTTC AGAGTGCTCA ATAGTAGTTA GAAAATGTTA GGTCATTTTA 550 CTTGTGCATT GTGATTCTTT TGGTTGTTGC TTACTGATCG ACGTGATGGA 600 TGGTTTACAG CTTCTTTCTG CTGTCAACTG CTGGTTTAAG TGTATGGTGC 650 TAGATOCTTA TAATCAARCA GACCATOOOG AGTGTAAGIC ACGGCCTGAC 700 NGTGGTCITT CAGCCATCAC AGAATTGGAT CCAGGTTACA TAACAGGTAG 750 TITCGGATIT TICITICTIT TGAGTTTTCT TCAATTTGAT ATCATCTTGT 800 TGTGATATAR TATGGCTARG TTCATTARTT TGGTCARTTT TCAGGTCCTC 250 TTTCTACTGT CTGGAAAGAG TGGCTTAAGT GGTGTGTTGA GTTTGGTATA 900 GAACCAMATG CAATTGTCGC TGTTCCATAC GATTGGAGAT TGTCACCAAC 950 CARRITGGAR DAGCGIGACC TITACTITCR CARGCICARG TIRGICCITA 1000 TOAGGCTART GTCTTTTATC TTCTCTTTTT ATGTARGATA AGCTARGAGC 1050 TCTGGTCGTC TICCTTTTTG CAGGTTGACC TTTGAAACTG CTTTAAAACT 1100 CCGIGGCGGC CCTTCTATAG TATTIGCCCA TICAATGGGT AATAATGTCT TONGATACTT TOTGGAATGG CTGAGGOTAG AANTTGCACO NAARCATTAT 1200 TTGAAGTGGC TTGATCAGCA TATCCATGCT TATTTCGCTG TTGGTACCGG 1250 CCTACTATCC TTANGTTACC ATTTTATTTT TTCTCTAATT GGGGGAGTTA TGTTGTGACT TACTGGATTG AGCTCGATAC CTGATTTGTT GTTGATTTAG 1350 GAGCTCCTCT TCTTGGTTCT GTTGAGGCAP. TCAAATCTAC TCTCTCTGGT 1400 GTANCETTIG GCCTTCCTGT TTCTGAGGTG ACCTCTGACT TCTCTTTAGT 1450 TTTAAGTAGT TGATATCAAC CAGGTCTTAT AACTCACTGG ATTTTCCTTT 1500 TGARAGTATT ACTITTGTTA ATTGARCTGC TGTACGCGAT ATGGTATCTG 1550 TAGATOTTGA AGTGCTAGTT ATCAAAGAAC ATATTGTGGG TAGTATACCT GTCAGCGGCC TTAGCTAATA CARCCAAACC ACATGTACAC TGATTTAGTT 1650 TTCAGATTAT TATGGTAGAC TTTAAGTTGA GAAGAAACTT TGACTGAAAT 1700 CITTITATTT TARTAGGCTA TGATTIGTTT ATTGARATCA TGTGACATAT 1750 TGACATGCGC TTCTCATGTT TTTTGTTGGC AAGGCTTCAG GGAACTGCTC 1800 GGTIGTTGTC CAATTCTTTT GCGTCGTCAT TGTGGCTTAT GCCATTTTCA 1350 AAGAATTGCA AGGGTGATAA CACATTCTGG ACGCATTTTT CTGGGGGTGC 1900 TGCNAAGARA GATARGCGCG TATACCACTG TGATGAAGAG GAATATCAAT 1950 CARACTATIC TGGCTGGCCG ACARATATTA TTARCATTGA ARTTCCTTCC ACTAGOGGTT AGACTOTGTA TATGOAACTG TAACACTAAC AAAAGTTTCA 2050 CCAAGAATGT TCACTCTCAT ATTTCGTTCC TTTGATGTGT ATCCATCAGT 2100 TACAGAARCA GCTCTAGTCA ACATGACCAG CATGGAATGT GGCCTTCCCA 2150 CCCTTTGTC TTTCACAGCC CGTGAACTAG CAGATGGGAC TCTTTTCAAA 2200 2250 GCNATAGAAG ACTATGACCC AGATAGCAAG AGGNTGTTAC ACCAGTTAAA GAAGTACOTA COTTTCTTTG TGATAAGAAA TATTGCTCAT CGATCATCAC TIGCIGGCIT CTIGTACGIC ARATIGITIT GITTARATCI CTATAICART 2350 IGITCATATG CTTTGTCTTT CTTACTATAA GAAACAAGTA TAATCAGAAA 2400 CCTTATTATT GATTATCAGT ICTCTCCTTA TATTATGGAA TGTCTTTTTC 2450 GITTACAGTT ATGARTGCAR ARGGGGGTAT TTTAGTTGAT TGATTCTCTC

ATTOTOTAGT TTGTTTTGAC TARTAGOGTO ARTTTTGTTT TTCTAGORAA TCTTTGTGAA TTATATATAA CATGCTAACT ATACTTTTCA GGTTGTATCA

TGATGACCCT GTTTTTAATC CTCTGACTCC TTGGGAGAGA CCACCTATAA AAAATGTATT TTGCATATAT GGTGCTCATC TAAAGACAGA GGTATGATGC

ATTOTCARTA TOACATTATG CGTTGACTTT GTTATTATAT TOCCCATTTG





TOTAL TOTAL	2800
GTTTGCANTA TCTTTTTGAA TTATGATTTA TCTTCTCCCT TGCATCTTAT	2850
COTATURACO GTTARAGGTA CIAMAIGIA: C.	2900
COTTATTACT TIGCCCCAAG IGGCAAACCI INICCIDATETT	2950
CACGGATATC ATTTATGAAA CIGAAGGTTC CCTCGTGTCA AGGIAATTTA	3000
CCCCALTGGC RGAAGTAAAA CAGGAAGGCA AAGICTA	3050
	3100
ANGINCTITY TIATCATICS, TITTGAGGTT AGTGGATGAT ANAGATACT	3150
PACTOCOLAG AGGTGTTGCA TGARACATGA CACITAGA	3200
	3250
CONCENTRAL BARRITGTT TTAAGAAACC GASTELL CONCENTRAL	2300
COMPARED TATCTGCAGG ICIGGAACIG IGGIACIG TOOLAGAATTAT	3350
CONTRACTOR COGREGACE GGTARGUICA GROUND TOTAL CACTO	3400
CONTROL A ACTACTGARG ACTAGONIAN TO THE CONTROL OF ACTACTTAT	3450
CONTROL TOT TOTAGTAC ACTUANTAL 10.000 CONTROL	3500
	3550
CONGRATTE SCIEGACCI RAGILLAGA	3600
CHARLETTERS CTTCCTCACC 1404000000000000000000000000000000000	3650
ACCOUNT TOT GTTGATTTAC CICUALITY IN THE CONTRACTOR TICG	3700
TOTAL COTTON GARC TIGIATIANT CIACACOTTO CATGOGRAC	3750
GARATARA CARCAGCCAG ARCACGATGG ARGCGREGA CATGACARAR GARATARA CARCAGCCAG ARCACGATGG CATAGCTAR CATGACARAR TARATGTTGA TCATGAGCAT GGGTCAGACA TCATAGCTAR AGAGCATTCC	3800
TANTETTEN TEATGAGEAT GGGTCAGACA TEATAGAGGATTCC	3850
GCACCAAGGG TTAAGTACA: ALL TAAAAAGTGGG TATTAA	3896
GGGGAAGAGA ACCGCAGTCT GGGAGCTTGA TAAAAGTGT	

<210> 11b <211> 709 <212> CDNA <213> COMA <400> 11

CTGGGGCCAA	AAGTGAACAT	AACAAGGACA	CCACAGTCAG	AGCATGATGT	so
TCAGATGTAC	AAGTGCATCT	AAATATAGAG	CATCAACATG	GTGAAGATAT.	j00
CATTCCCAAT	ATGACAAAGT	TACCTACAAT	GAAGTACATA	ACCTATTATG	150
AGGATTCTCA	AAGTTTTCCA	GGGACAAGAA	CAGCAGTTTG	GGAGCTTGAT	200
AARGCAAATC	ACAGGAACAT	TGTCAGATCT	CCAGCTTTGA	TGCGGGAGCT	250
GTGGCTTGAG	ATGTGGCATG	ATATTCATCC	TGATAAAAG	TCCAAGTTTG	300
TTACAAAAGG	TGGTGTCTGA	TCCTCACTAT	TTTCTTCTAT	AAATGTTTGA	350
GTITGTATTG	ACATTGTAAG	TATTGCAACA	AAAAGCAAAG	CGTGGGCCTC	400
TGAGGGNTGA	GGACTGCTAT	TGGGATTACG	GGAAAGCTCG	ATGTGCATGG	450
GCTGAACATT	GTGAATACAG	GTTAGAATAT	TCALATTATA	TTTTGCAAAA	500
TATTCTCTTT	TIGTGTATTI	AGGCCACCTT	TCCCCGGTCA	CAACGATGCA	550
GATATGTATT	CGGGGATGTI	CACCIGGGAC	: AGAGTTGCAG	ATTGAAGAGT	600
TOTACATOTO	ACATCCTGTO	ACACTATGTO	GATATTTA	. GAAACTTTGT	€50
TTGGCGGAAC	AACAAGTTT	CACAAACATI	TGAAGAAGA	AGCGAAATGA	700
TICAGAGAG					709

THIS PAGE BLANK (USPTO)



### (12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

# (19) World Intellectual Property Organization International Bureau





### (43) International Publication Date 12 October 2000 (12.10.2000)

PCT

# (10) International Publication Number WO 00/60095 A3

(51) International Patent Classification<sup>7</sup>: C12N 15/54, 9/10, 15/81, 15/82, 1/16, 5/10, A01K 67/027, C12P 7/64

(21) International Application Number: PCT/EP00/02701

(22) International Filing Date: 28 March 2000 (28.03.2000)

(25) Filing Language: English

(26) Publication Language:

English

(30) Priority Data:

99106656.4 1 April 1999 (01.04.1999) EP 99111321.8 10 June 1999 (10.06.1999) EP 60/180,687 7 February 2000 (07.02.2000) US

(71) Applicant (for all designated States except US): BASF PLANT SCIENCE GMBH [DE/DE]; D-67056 Ludwigshafen (DE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): DAHLQVIST, Anders [SE/SE]; Hemmansvägen 2, S-244 66 Furulund (SE). STAHL, Ulf [SE/SE]; Liljegatan 7b, S-753 24 Uppsala (SE). LENMAN, Marit [SE/SE]; Revingegatan 13a, S-223 59 Lund (SE). BANAS, Antoni [SE/PL]; Wiolinowa 14, PL-08110 Siedlee (PL). RONNE, Hans [SE/SE]; Dirigentvägen 169, S-756 54 Uppsala (SE). STYMNE, Sten [SE/SE]; Torrlösa 1380, S-269 90 Svalöv (SE).

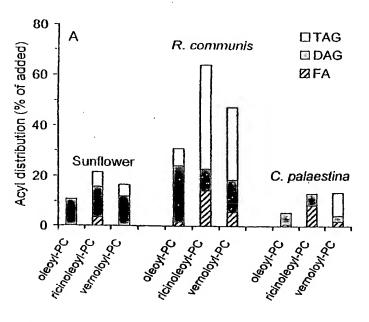
- (74) Agent: FITZNER, Uwe; Lintorfer Str. 10, D-40878 Ratingen (DE).
- (81) Designated States (national): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

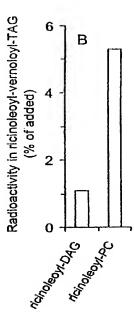
#### Published:

With international search report.

[Continued on next page]

(54) Title: ENZYMES OF THE BIOSYNTHETIC PATHWAY FOR THE PRODUCTION OF TRIACYLGLYCEROL AND RE-COMBINANT DNA MOLECULES ENCODING THESE ENZYMES





(57) Abstract: The present invention relates to the isolation, identification and characterization of nucleotide sequences encoding an enzyme catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol, to the said enzymes and a process for the production of triacylglycerols.

SDOCID: <WO 0060095A3 1 3





(88) Date of publication of the international search report: 1 February 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

a. classification of subject matter IPC 7 C12N15/54 C12N9/10

C12N5/10

A01K67/027

C12N15/81 C12P7/64

C12N15/82

C12N1/16

According to International Patent Classification (IPC) or to both national classification and IPC

#### B. FIELDS SEARCHED

 $\begin{array}{ll} \mbox{Minimum documentation searched (classification system followed by classification symbols)} \\ \mbox{IPC 7} & \mbox{C12N} & \mbox{A01K} & \mbox{C12P} \\ \end{array}$ 

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic state base consulted during the international search (name of data base and, where practical, search terms used)

STRAND, EPO-Internal, WPI Data, MEDLINE, CHEM ABS Data, BIOSIS, EMBL

Category *	Citation of document, with indication, where appropriate, of	the relevant passages	Relevant to daim No.
X	PETER VERHASSELT ET AL.: "Tw reading frames revealed in th segent flanking the centromer Saccharomyces cerevisiae chro right arm" YEAST, vol. 10, no. 7, July 1994 (19 1355-1361, XP002112572	e 23.6kb e on the mosome XIV	1-23,27
X	abstract; table 2 -& Swissprot Database Entry Y Accession number P40345; 1 Fe XP002112574 the whole document		1-23,27
		-/	
X Furt	ther documents are listed in the continuation of box C.	Patent family members are listed	1 in annex.
"A" docum consi "E" earlier filing "L" docum which citatic "O" docum other "P" docum	ategones of cited documents:  tent defining the general state of the art which is not dered to be of particular relevance document but published on or after the international date ent which may throw doubts on priority claim(s) or n is cited to establish the publication date of another on or other special reason (as specified) tent referring to an oral disclosure, use, exhibition or means the published prior to the international filing date but than the priority date claimed	"T" later document published after the int or priority date and not in conflict with cited to understand the principle or the invention.  "X" document of particular relevance; the cannot be considered novel or cannot involve an inventive step when the decument of particular relevance; the cannot be considered to involve an indocument is combined with one or ments, such combination being obvirin the art.  "&" document member of the same paten	n the application but neory underlying the claimed invention of the considered to comment is taken alone claimed invention inventive step when the lore other such docupous to a person skilled
Date of the	actual completion of the international search	Date of mailing of the international se	
	17 October 2000	30/10/2000	
1	17 October 2000		



Interna Spilication No PCT/En 0/02701

ation) DOCUMENTS CONSIDERED TO BE RELEVANT	
	Relevant to claim No.
DATABASE EMBL 'Online! Database Entry SPBC776, 21 January 1999 (1999-01-21) LYNE M. ET AL.: "S. pombe chromosome II cosmid c776" Database accession no. AL035263 XP002150203 the whole document	1-23,27
DATABASE EMBL 'Online! Database Entry AI398644, 10 February 1999 (1999-02-10) XP002150204 the whole document & MARY ANNE NELSON ET AL.: "Expressed sequences from conidial, mycelial, and sexual stages of Neurospora crassa "FUNGAL GENETICS AND BIOLOGY, vol. 21, 1997, pages 348-363, XP000952173	1-23,27
KEITH STOBART ET AL.: "Triacylglycerols are synthesized and utilized by transacylation reactions in microsomal preparations of developing safflower (Carthamus tinctorius L.) seeds" PLANTA, vol. 203, no. 1, 1997, pages 58-66, XP002112573 page 58, right-hand column, last paragraph -page 59, left-hand column, paragraph 1 page 63, right-hand column, paragraph 2	25
WO 98 55631 A (CALGENE LLC) 10 December 1998 (1998-12-10) page 9, line 36 -page 10, line 7 page 12, line 28 -page 13, line 18 page 14, line 34 -page 15, line 13 page 20, line 5 -page 25, line 4	1-27
DATABASE SWALL 'Online! Database Entry 094680, 1 May 1999 (1999-05-01) LYNE M. ET AL.: "hypothetical 69.7 kDa protein C776.14 in chromosome II" Database accession no. 094680 XP002150205 the whole document	1-23,27
	DATABASE EMBL 'Online! Database Entry SPBC776, 21 January 1999 (1999-01-21) LYNE M. ET AL.: "S. pombe chromosome II cosmid c776" Database accession no. AL035263 XP002150203 the whole document  DATABASE EMBL 'Online! Database Entry A1398644, 10 February 1999 (1999-02-10) XP002150204 the whole document & MARY ANNE NELSON ET AL.: "Expressed sequences from conidial, mycelial, and sexual stages of Neurospora crassa " FUNGAL GENETICS AND BIOLOGY, vol. 21, 1997, pages 348-363, XP000952173  KEITH STOBART ET AL.: "Triacylglycerols are synthesized and utilized by transacylation reactions in microsomal preparations of developing safflower (Carthamus tinctorius L.) seeds" PLANTA, vol. 203, no. 1, 1997, pages 58-66, XP002112573 page 58, right-hand column, last paragraph -page 59, left-hand column, paragraph 1 page 63, right-hand column, paragraph 1 page 63, right-hand column, paragraph 2  WO 98 55631 A (CALGENE LLC) 10 December 1998 (1998-12-10) page 9, line 36 -page 10, line 7 page 12, line 28 -page 13, line 18 page 14, line 34 -page 15, line 13 page 20, line 5 -page 25, line 4  DATABASE SWALL 'Online! Database Entry 094680, 1 May 1999 (1999-05-01) LYNE M. ET AL.: "hypothetical 69.7 kDa protein C776.14 in chromosome II" Database accession no. 094680 XP002150205

2

		MAL SEARCE		11	Interna 'al	00/02701
Patent document cited in search report	nt	Publication date	P	atent family member(s)	,	Publication date
WO 9855631	Α	10-12-1998	CN EP	12664 10038	60 T 882 A	13-09-2000 31-05-2000

THIS PAGE BLANK (USPTO)

# This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

□ BLACK BORDERS
□ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
□ FADED TEXT OR DRAWING
□ BLURRED OR ILLEGIBLE TEXT OR DRAWING
□ SKEWED/SLANTED IMAGES
□ COLOR OR BLACK AND WHITE PHOTOGRAPHS
□ GRAY SCALE DOCUMENTS
□ LINES OR MARKS ON ORIGINAL DOCUMENT
□ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

# IMAGES ARE BEST AVAILABLE COPY.

□ OTHER: \_\_\_\_\_

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

# THIS PAGE BLANK (USPT!)